

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:50:06 ; Search time 13 Seconds  
(Without alignments)  
230,896 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 49  
Sequence: 1 YLYQWLGAPVPYPPDLXPRR.....DELADHIGFOEAYRRFGPV 49

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 193982 seqs, 61258239 residues

Word size : 0

Total number of hits satisfying chosen parameters: 193982

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCN\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	25	51.0	58	US-09-724-676-90318	Sequence 90318, A
2	25	51.0	58	US-09-724-676A-90318	Sequence 90318, A
3	25	51.0	100	US-10-283-656-1	Sequence 1, Appl
4	10	20.4	13	US-09-802-154-21	Sequence 21, Appl
5	10	20.4	257	US-10-096-246-2	Sequence 2, Appl
6	7	14.3	433	US-10-278-945-2	Sequence 2, Appl
7	6	12.2	220	PCT-US02-32727-14821	Sequence 14821, A
8	6	12.2	451	US-10-258-662-13	Sequence 13, Appl
9	6	12.2	474	PCT-US02-32727-8047	Sequence 8047, Ap
10	6	12.2	509	PCT-US02-32727-29930	Sequence 29930, A
11	6	12.2	515	PCT-US02-32727-21198	Sequence 21198, A
12	6	12.2	600	PCT-US02-32727-25938	Sequence 25938, A
13	6	12.2	694	US-09-724-676-63352	Sequence 63352, A
14	6	12.2	694	US-09-724-676-63353	Sequence 63353, A
15	6	12.2	694	US-09-724-676A-63352	Sequence 63352, A
16	6	12.2	694	US-09-724-676A-63353	Sequence 63353, A
17	6	12.2	902	US-09-724-676-63356	Sequence 63356, A
18	6	12.2	902	US-09-724-676A-63356	Sequence 63356, A
19	6	12.2	934	PCT-US02-32637-31	Sequence 31, Appl
20	6	12.2	934	US-10-270-839-31	Sequence 31, Appl
21	6	12.2	957	US-09-724-676-63354	Sequence 63354, A
22	6	12.2	957	US-09-724-676A-63354	Sequence 63354, A
23	6	12.2	958	US-09-724-676-87329	Sequence 87329, A
24	6	12.2	958	US-09-724-676A-87329	Sequence 87329, A
25	6	12.2	1021	US-09-724-676-87327	Sequence 87327, A
26	6	12.2	1021	US-09-724-676A-87327	Sequence 87327, A

27	6	12.2	1183	US-09-724-676-87328	Sequence 87328, A
28	6	12.2	1183	US-09-724-676A-87328	Sequence 87328, A
29	6	12.2	1484	US-09-724-676-79242	Sequence 79242, A
30	6	12.2	1484	US-09-724-676-79252	Sequence 79252, A
31	6	12.2	1484	US-09-724-676A-79242	Sequence 79242, A
32	6	12.2	1484	US-09-724-676A-79252	Sequence 79252, A
33	6	12.2	1493	US-09-724-676-79245	Sequence 79245, A
34	6	12.2	1493	US-09-724-676-79255	Sequence 79255, A
35	6	12.2	1493	US-09-724-676A-79245	Sequence 79245, A
36	6	12.2	1493	US-09-724-676A-79255	Sequence 79255, A
37	6	12.2	1836	US-09-724-676-79241	Sequence 79241, A
38	6	12.2	1836	US-09-724-676-79251	Sequence 79251, A
39	6	12.2	1836	US-09-724-676A-79241	Sequence 79241, A
40	6	12.2	1836	US-09-724-676A-79251	Sequence 79251, A
41	6	12.2	1845	US-09-724-676-79244	Sequence 79244, A
42	6	12.2	1845	US-09-724-676-79254	Sequence 79254, A
43	6	12.2	1845	US-09-724-676A-79244	Sequence 79244, A
44	6	12.2	1845	US-09-724-676A-79254	Sequence 79254, A
45	6	12.2	1903	US-09-724-676-79239	Sequence 79239, A

## ALIGNMENTS

```
RESULT 1
US-09-724-676-90318
; Sequence 90318, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90318
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-90318

Query Match          51.0%; Score 25; DB 5; Length 58;
Best Local Similarity 100.0%; Pred. No. 6; 2e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 34 LNPDCDELADHIGFOEAYRRFGPV 58

RESULT 2
US-09-724-676A-90318
; Sequence 90318, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90318
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-90318

Query Match          51.0%; Score 25; DB 5; Length 58;
Best Local Similarity 100.0%; Pred. No. 6; 2e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 34 LNPDCDELADHIGFOEAYRRFGPV 58
```

Db 34 LNPDCDELADHIGFOEAYRRFGPV 58

RESULT 3  
US-10-283-656-1  
; Sequence 1, Application US/10283656  
; GENERAL INFORMATION:  
; APPLICANT: EKEMA, George Mbella  
; APPLICANT: MAYS, Robert W.  
; APPLICANT: BRUNDEN, Kurt R.  
; TITLE OF INVENTION: Methods for Using Osteocalcin  
; FILE REFERENCE: ATX-005  
; CURRENT APPLICATION NUMBER: US/10/283,656  
; CURRENT FILING DATE: 2002-10-29  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 100  
; TYPE: PR  
; ORGANISM: Homo sapiens  
US-10-283-656-1

Query Match 51.0%; Score 25; DB 6; Length 100;  
Best Local Similarity 100.0%; Pred. No. 9,9e-20;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
Db 76 LNPDCDELADHIGFOEAYRRFGPV 100

RESULT 4  
US-09-802-154-21  
; Sequence 21, Application US/09802154  
; GENERAL INFORMATION:  
; APPLICANT: Itoh, Nobuyuki  
; APPLICANT: Kavanaugh, W. Michael  
; TITLE OF INVENTION: HUMAN FGF-23 GENE AND GENE EXPRESSION  
; FILE REFERENCE: PP-17149.001/201130.409  
; CURRENT APPLICATION NUMBER: US/09/802,154  
; CURRENT FILING DATE: 2001-03-07  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 13  
; TYPE: PR  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: E tag  
US-09-802-154-21

Query Match 20.4%; Score 10; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDPL 16  
Db 1 GAVPYPPDPL 10

RESULT 5  
US-10-096-246-2  
; Sequence 2, Application US/10096246  
; GENERAL INFORMATION:  
; APPLICANT: The Minister of National Defence, Government of Canada  
; APPLICANT: Fulton, R E  
; APPLICANT: Alvi, Azhar E  
; TITLE OF INVENTION: Cloning, Expression, Sequencing, and Functional Enhancement of M  
; FILE REFERENCE: scrv Antibody Against Venezuelan Equine Encephalitis Virus (Vee)  
; CURRENT APPLICATION NUMBER: US/10/096,246

; CURRENT FILING DATE: 2002-03-13  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 257  
; TYPE: PR  
; ORGANISM: Mouse hybridoma cell line 1A4A1  
US-10-096-246-2

Query Match 20.4%; Score 10; DB 6; Length 257;  
Best Local Similarity 100.0%; Pred. No. 0.0024;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDPL 16  
Db 243 GAVPYPPDPL 252

RESULT 6  
US-10-278-945-2  
; Sequence 2, Application US/10278945  
; GENERAL INFORMATION:  
; APPLICANT: Cao et al.  
; TITLE OF INVENTION: Retinoic Acid Receptor Epsilon  
; FILE REFERENCE: PF123D1C1  
; CURRENT APPLICATION NUMBER: US/10/278,945  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: US 09/022,789  
; PRIOR FILING DATE: 1998-02-12  
; PRIOR APPLICATION NUMBER: US 08/466,120  
; PRIOR FILING DATE: 1995-06-06  
; PRIOR APPLICATION NUMBER: PCT/US94/07266  
; PRIOR FILING DATE: 1994-06-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 433  
; TYPE: PR  
; ORGANISM: Homo sapiens  
US-10-278-945-2

Query Match 14.3%; Score 7; DB 6; Length 433;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WLGAPVP 11  
Db 4 WLGAPVP 10

RESULT 7  
PCT-US02-32727-14821  
; Sequence 14821, Application PC/TUS0232727  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skelky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Sigling  
; APPLICANT: Jen, Shyian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darrick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglass, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; CURRENT FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992

```
; SEQ ID NO 14821
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Propionl acnes
PCT-US02-32727-14821
```

```
Query Match          12.2%; Score 6; DB 1; Length 220;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 LGAPVP 11
      |||||
Db 166 LGAPVP 171
```

```
RESULT 8
US-10-258-662-13
; Sequence 13, Application US/10258662
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preetl
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YAO, Monique G.
; APPLICANT: BURFORD, Neil
; APPLICANT: BATRA, Sajeev
; APPLICANT: POLICKY, Jennifer J.
; TITLE OF INVENTION: RNA METABOLISM PROTEINS
; FILE REFERENCE: PF-0771 US#
; CURRENT APPLICATION NUMBER: US/10/258,662
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/201,875
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/200,184
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,090
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/210,232
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 60/220,553
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 2472784CD1
US-10-258-662-13

Query Match          12.2%; Score 6; DB 6; Length 451;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 LGAPVP 11
      |||||
Db 49 LGAPVP 54
```

```
RESULT 9
PCT-US02-32727-8047
; Sequence 8047, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
```

```
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siding
; APPLICANT: Jen, Shylian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 8047
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Propionl acnes
PCT-US02-32727-8047
```

```
Query Match          12.2%; Score 6; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 39 QEAYRR 44
      |||||
Db 365 QEAYRR 370
```

```
RESULT 10
PCT-US02-32727-29930
; Sequence 29930, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siding
; APPLICANT: Jen, Shylian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 29930
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Propionl bacterium acnes
PCT-US02-32727-29930

Query Match          12.2%; Score 6; DB 1; Length 509;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 39 QEAYRR 44
      |||||
Db 400 QEAYRR 405
```

```
RESULT 11
PCT-US02-32727-21198
; Sequence 21198, Application PC/TUS0232727
; GENERAL INFORMATION:
```

APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shylan  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darlick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 21198  
LENGTH: 515  
TYPE: PRT  
ORGANISM: Propionl acnes  
PCT-US02-32727-21198

Query Match 12.2%; Score 6; DB 1; Length 515;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 APVPY 13  
DB 210 APVPY 215

RESULT 12  
PCT-US02-32727-25938  
Sequence 25938, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shylan  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darlick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 25938  
LENGTH: 600  
TYPE: PRT  
ORGANISM: Propionl acnes  
PCT-US02-32727-25938

Query Match 12.2%; Score 6; DB 1; Length 600;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 DCDELA 33  
DB 32 DCDELA 37

RESULT 13

US-09-724-676-63352  
Sequence 63352, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724.676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 63352  
LENGTH: 694  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676-63352

Query Match 12.2%; Score 6; DB 5; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
DB 19 LGAPVP 24

RESULT 14  
US-09-724-676-63353  
Sequence 63353, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724.676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 63353  
LENGTH: 694  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676-63353

Query Match 12.2%; Score 6; DB 5; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
DB 19 LGAPVP 24

RESULT 15  
US-09-724-676A-63352  
Sequence 63352, Application US/09724676A  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724.676A  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 63352  
LENGTH: 694  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-63352

Query Match 12.2%; Score 6; DB 5; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11



Db 19 LGAPVP 24

RESULT 16  
US-09-724-676A-63353

; Sequence 63353, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 63353  
; LENGTH: 694  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-63353

Query Match 12.2%; Score 6; DB 5; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
Db 19 LGAPVP 24

RESULT 17  
US-09-724-676-63356

; Sequence 63356, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 63356  
; LENGTH: 902  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-63356

Query Match 12.2%; Score 6; DB 5; Length 902;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
Db 227 LGAPVP 232

RESULT 18  
US-09-724-676A-63356

; Sequence 63356, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 63356  
; LENGTH: 902  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-63356

Query Match 12.2%; Score 6; DB 5; Length 902;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
Db 227 LGAPVP 232

RESULT 19  
PCT-US02-32637-31

; Sequence 31, Application PC/TUS0232637  
; GENERAL INFORMATION:  
; APPLICANT: Morphotek, Inc.  
; TITLE OF INVENTION: Genetic Hypermutability of Plants for Gene Discovery and Diagn  
; FILE REFERENCE: AG0002PCT (MOR-0137)  
; CURRENT APPLICATION NUMBER: PCT/US02/32637  
; CURRENT FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/328,750  
; PRIOR FILING DATE: 2001-10-12  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 934  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-32637-31

Query Match 12.2%; Score 6; DB 1; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12  
Db 614 GAPVPY 619

RESULT 20  
US-10-270-839-31

; Sequence 31, Application US/10270839  
; GENERAL INFORMATION:  
; APPLICANT: Chao, Qimin  
; APPLICANT: Grasso, Luigi  
; APPLICANT: Sass, Philip M.  
; TITLE OF INVENTION: Genetic Hypermutability of Plants for Gene Discovery and Diagn  
; FILE REFERENCE: AG0002US (MOR-0133)  
; CURRENT APPLICATION NUMBER: US/10/270, 839  
; CURRENT FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/328,750  
; PRIOR FILING DATE: 2001-10-12  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 934  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-270-839-31

Query Match 12.2%; Score 6; DB 6; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12  
Db 614 GAPVPY 619

RESULT 21  
US-09-724-676-63354  
; Sequence 63354, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD

```
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63354
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-63354
```

```
Query Match          12.2%; Score 6; DB 5; Length 957;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 LGAPVP 11
|111111
DB 282 LGAPVP 287
```

```
RESULT 22
US-09-724-676A-63354
; Sequence 63354, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63354
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-63354
```

```
Query Match          12.2%; Score 6; DB 5; Length 957;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 LGAPVP 11
|111111
DB 282 LGAPVP 287
```

```
RESULT 23
US-09-724-676-87329
; Sequence 87329, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87329
; LENGTH: 958
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-87329
```

```
Query Match          12.2%; Score 6; DB 5; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 LGAPVP 11
|111111
DB 94 LGAPVP 99
```

```
RESULT 24
US-09-724-676A-87329
; Sequence 87329, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87329
; LENGTH: 958
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-87329
```

```
Query Match          12.2%; Score 6; DB 5; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 LGAPVP 11
|111111
DB 94 LGAPVP 99
```

```
RESULT 25
US-09-724-676-87327
; Sequence 87327, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87327
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-87327
```

```
Query Match          12.2%; Score 6; DB 5; Length 1021;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 LGAPVP 11
|111111
DB 94 LGAPVP 99
```

```
RESULT 26
US-09-724-676A-87327
; Sequence 87327, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87327
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-87327
```

```
Query Match          12.2%; Score 6; DB 5; Length 1021;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 6 LGAPVP 11  
|11111  
Db 94 LGAPVP 99

RESULT 27  
US-09-724-676-87328

; Sequence 87328, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 87328  
; LENGTH: 1183  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-87328

Query Match 12.2%; Score 6; DB 5; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11  
|11111  
Db 94 LGAPVP 99

RESULT 28  
US-09-724-676A-87328

; Sequence 87328, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 87328  
; LENGTH: 1183  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-87328

Query Match 12.2%; Score 6; DB 5; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11  
|11111  
Db 94 LGAPVP 99

RESULT 29  
US-09-724-676-79242

; Sequence 79242, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 79242  
; LENGTH: 1484  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-79242

Query Match 12.2%; Score 6; DB 5; Length 1484;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36  
|11111  
Db 1332 ELADHI 1337

RESULT 30  
US-09-724-676-79252

; Sequence 79252, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 79252  
; LENGTH: 1484  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-79252

Query Match 12.2%; Score 6; DB 5; Length 1484;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36  
|11111  
Db 1332 ELADHI 1337

RESULT 31  
US-09-724-676A-79242

; Sequence 79242, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 79242  
; LENGTH: 1484  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-79242

Query Match 12.2%; Score 6; DB 5; Length 1484;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36  
|11111  
Db 1332 ELADHI 1337

RESULT 32  
US-09-724-676A-79252

; Sequence 79252, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2

```

; SEQ ID NO 79252
; LENGTH: 1484
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-724-676A-79252

```

Query Match	12.2%;	Score 6;	DB 5;	Length 1484;
Best Local Similarity	100.0%;	Pred. No. 2e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	31	ELADHI	36
Db	1332	ELADHI	1337

```

RESULT 33
US-09-724-676-79245
; Sequence 79245; Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentln version 3.2
; SEQ ID NO 79245
; LENGTH: 1493
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-724-676-79245

```

Query Match	12.2%;	Score 6;	DB 5;	Length 1493;
Best Local Similarity	100.0%;	Pred. No. 2e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	31	ELADHI	36
	11111		
Db	1341	ELADHI	1346

RESULT 34  
 US-09-724-676-79256  
 ; Sequence 79256, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724, 676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: Patentln version 3.2  
 ; SEQ ID NO 79256  
 ; LENGTH: 1493  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-79256

Query Match	12.2%;	Score 6;	DB 5;	Length 1493;
Best Local Similarity	100.0%;	Pred. No. 2e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	31	ELADHI	36
Db	1341	ELADHI	1346

RESULT 35  
US-09-724-676A-79245  
; Sequence 79245, Application US/09724676A  
; GENERAL INFORMATION:  
? ; APPLICANT: Comugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing

```

: FILE REFERENCE: 129181.4 Compusen
: CURRENT APPLICATION NUMBER: US/09/724,676A
: CURRENT FILING DATE: 2000-11-28
: NUMBER OF SEQ. ID NOS: 97222
: SOFTWARE: PatentIn version 3.2
: SEQ. ID NO 79245

```

ORGANISM: Homo sapiens  
US-09-724-676A-79245

Query Match	12.2%	Score 6;	DB 5;	Length 1493;
Best Local Similarity	100.0%	Pred. No. 2e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	31	ELADHI	36
Db	1341	ELADHI	1346

```

RESULT 36
US-09-724-676A-79256
; Sequence 79256, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ. ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79256
; LENGTH: 1493
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-79256

```

Query Match	12.2%;	Score 6;	DB 5;	Length 1493;
Best Local Similarity	100.0%;	Pred. No. 2e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	31	ELADHI	36
Db	1341	ELADHI	1346

```

RESULT 37
US-09-724-676-79241
; Sequence 79241, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ. ID NOS. 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79241
; LENGTH: 1836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-79241

```

Query Match	12.2%;	Score 6;	DB 5;	Length 1836;
Best Local Similarity	100.0%;	Pred. No. 2.4e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	31	ELADHI	36
Db	1332	ELADHI	1337

### RESULT 38

```
US-09-724-676-79251
; Sequence 79251, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79251
; LENGTH: 1836
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-79251

Query Match          12.2%; Score 6; DB 5; Length 1836;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36
Db 1332 ELADHI 1337

RESULT 39
US-09-724-676A-79241
; Sequence 79241, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79241
; LENGTH: 1836
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-79241

Query Match          12.2%; Score 6; DB 5; Length 1836;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36
Db 1332 ELADHI 1337

RESULT 40
US-09-724-676A-79251
; Sequence 79251, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79251
; LENGTH: 1836
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-79251

Query Match          12.2%; Score 6; DB 5; Length 1836;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36
Db 1332 ELADHI 1337
```

```
Db 1332 ELADHI 1337

RESULT 41
US-09-724-676-79244
; Sequence 79244, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79244
; LENGTH: 1845
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-79244

Query Match          12.2%; Score 6; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36
Db 1341 ELADHI 1346

RESULT 42
US-09-724-676-79255
; Sequence 79255, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79255
; LENGTH: 1845
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-79255

Query Match          12.2%; Score 6; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36
Db 1341 ELADHI 1346

RESULT 43
US-09-724-676A-79244
; Sequence 79244, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79244
; LENGTH: 1845
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-79244
```

Query Match 12.2%; Score 6; DB 5; Length 1845;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
|||||  
Db 1341 ELADHI 1346

## RESULT 44

US-09-724-676A-79255  
; Sequence 79255, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724.676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 79255  
; LENGTH: 1845  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-79255

Query Match 12.2%; Score 6; DB 5; Length 1845;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
|||||  
Db 1341 ELADHI 1346

## RESULT 45

US-09-724-676-79239  
; Sequence 79239, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724.676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 79239  
; LENGTH: 1903  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-79239

Query Match 12.2%; Score 6; DB 5; Length 1903;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
|||||  
Db 1332 ELADHI 1337

Search completed: December 4, 2002, 15:54:12  
Job time : 15 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:49:06 : Search time 19 Seconds  
(without alignments)  
247.926 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 49

Sequence: 1 YLYQWLGAAPYPPDPLKPRR.....DELADHIGFOEAYRRFGPV 49

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	25	51.0	48	2	S02208 osteocalcin - emu
2	25	51.0	49	1	GEMKI osteocalcin - crab
3	25	51.0	100	1	GEHU osteocalcin precur
4	25	51.0	100	1	GEBO osteocalcin precur
5	15	30.6	49	1	GECT osteocalcin - cat
6	14	28.6	97	1	GRCH osteocalcin precur
7	11	22.4	99	1	GBRT osteocalcin precur
8	10	20.4	49	1	A61280 osteocalcin - rabb
9	8	16.3	536	2	T42606 probable transcrip
10	7	14.3	275	2	E97419 hypotetical prote
11	7	14.3	286	2	B84226 3-hydroxyacyl-CoA
12	7	14.3	317	2	G83544 probable transcrip
13	7	14.3	447	2	I38975 nuclear orphan rec
14	7	14.3	586	1	WMBEDE 65K early nonstruc
15	7	14.3	587	1	WMBERE 65K early nonstruc
16	7	14.3	4767	2	T31345 hypotetical prote
17	6	12.2	70	2	C71923 ribosomal protein
18	6	12.2	70	2	B64590 L71-1 protein - fr
19	6	12.2	100	2	S62333 yidd family [impor
20	6	12.2	103	2	E86364 yidd family - Chla
21	6	12.2	103	2	H72060 conserved hypotet
22	6	12.2	106	2	E81606 myohemerithrin - p
23	6	12.2	119	2	S38261 DEAD box protein a
24	6	12.2	136	2	T18052 cystatin precursor
25	6	12.2	139	1	UDCH hypotetical prote
26	6	12.2	143	2	T12144 Zn-alpha 2-glycopr
27	6	12.2	164	2	PC4154 hypotetical prote
28	6	12.2	169	2	S03744 C2C2.6 protein -
29	6	12.2	171	2	S44737

30	6	12.2	185	2	AH1075 probable fibrillar
31	6	12.2	214	2	AB2392 hypotetical prote
32	6	12.2	235	2	DB1201 cell division prot
33	6	12.2	242	2	G81776 acetyltransferase
34	6	12.2	252	2	AH3267 probable transcrip
35	6	12.2	264	2	T35168 hypotetical prote
36	6	12.2	289	2	A87646 quinolinate synthe
37	6	12.2	304	2	C69111 conserved hypotet
38	6	12.2	332	2	E75473 NADH-dependent fla
39	6	12.2	338	2	E69964 hypotetical prote
40	6	12.2	352	2	C70532 hypotetical prote
41	6	12.2	367	2	E64399 hypotetical prote
42	6	12.2	370	2	S32173 hypotetical prote
43	6	12.2	397	2	G70796 hypotetical prote
44	6	12.2	399	2	A33396 beta-N-acetylgluco
45	6	12.2	420	2	T06000 aspartic proteinase

## ALIGNMENTS

```
RESULT 1
S02208
osteocalcin - emu
C:Species: Dromaius novaehollandiae (emu)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 18-Jun-1993
C:Accession: S02208
R:Hug, N.L.; Tseng, A.; Chapman, G.E.
Biochem. Int. 15, 271-277, 1987
A:Title: The amino acid sequence of emu osteocalcin: gas phase sequencing of Gla-cont
A:Reference number: S02208; MUID:88134266; PMID:3501719
A:Accession: S02208
A:Molecule type: protein
A:Residues: 1-48 <HUQ>
C:Superfamily: osteocalcin

Query Match          51.0%: Score 25; DB 2; Length 48;
Best Local Similarity 100.0%: Pred. No. 1.9e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 24 LNPDCDELADHIGFOEAYRRFGPV 48

RESULT 2
GEMKI
osteocalcin - crab-eating macaque
N:Alternate names: BGP: bone Gla protein; gamma-carboxyglutamic acid-containing prote
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 06-Sep-1996
C:Accession: A03302
R:Hauschka, P.V.; Carr, S.A.; Blemann, K.
Biochemistry 21, 638-642, 1982
A:Title: Primary structure of monkey osteocalcin.
A:Reference number: A03302; MUID:82182842; PMID:6978733
A:Accession: A03302
A:Molecule type: protein
A:Residues: 1-49 <HAU>
C:Comment: This protein, isolated from bone, binds strongly to apatite.
C:Superfamily: osteocalcin
C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline
F:9/Modified site: 4-hydroxyproline (Pro) #status experimental
F:17,21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental
F:23-29/Disulfide bonds: #status experimental

Query Match          51.0%: Score 25; DB 1; Length 49;
Best Local Similarity 100.0%: Pred. No. 1.9e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 25 LNPDCDELADHIGFOEAYRRFGPV 49
```

RESULT 3  
GENU

osteocalcin precursor [validated] - human  
N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing protein  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1980 #sequence\_revision 07-Apr-1994 #text\_change 08-Dec-2000  
C:Accession: S12652; C25471; A03301; S08694  
R:Kiefer, M.C.; Saphire, A.C.S.; Bauer, D.M.; Barr, P.J.  
Nucleic Acids Res. 18, 1909, 1990  
A:title: The cDNA and derived amino acid sequences of human and bovine bone Gla protein.  
A:Reference number: S12652; M01D:90245603; PMID:2336375  
A:Accession: S12652  
A:Molecule type: mRNA  
A:Residues: 1-100 <RTE>  
A:Cross-references: EMBL:X53698; NID:936092; PIDN:CAA37736.1; PID:936093  
R:Gleeson, A.J.; Rosen, V.; Buckner, J.L.; Kite, R.; Wang, E.A.; Wozney, J.M.  
EMBO J. 5, 1885-1890, 1986  
A:title: Isolation of the human gene for bone gla protein utilizing mouse and rat cDNA  
A:Reference number: A91045; M01D:87004555; PMID:3019668  
A:Accession: C25471  
A:Molecule type: DNA  
A:Residues: 1-32,35-100 <CELE>  
A:Cross-references: EMBL:X04143; NID:929449; PIDN:CAA27763.1; PID:929450  
R:Poser, J.W.; Esch, F.S.; Ling, N.C.; Price, P.A.  
J. Biol. Chem. 255, 8685-8691, 1980  
A:title: Isolation and sequence of the vitamin K-dependent protein from human bone. Unde  
A:Reference number: A03301; M01D:81006914; PMID:6967872  
A:Accession: A03301  
A:Molecule type: protein  
A:Residues: 52-100 <POS>  
R:Cairns, J.R.; Williamson, M.K.; Price, P.A.  
Anal. Biochem. 199, 93-97, 1991  
A:title: Direct identification of gamma-carboxyglutamic acid in the sequencing of vitamin  
A:Reference number: A44566; M01D:9222218; PMID:1807167  
A:Contents: annotation  
C:Comment: This protein, isolated from bone, binds strongly to apatite.  
C:Comment: Alternative splicing may produce the sequence presented in reference A91045.  
C:Comment: Glu-68 is gamma-carboxylated in 9-50% of the molecules.  
C:Genetics:  
A:Gene: GDB:BGLAP  
A:Cross-references: GDB:118760; OMIM:112260  
A:Map position: 1q25-1q31  
A:Insertions: 22/1; 35/1; 58/2  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid; extracellular matrix  
F:1-51/Domain: signal sequence #status predicted <SIG>  
F:52-100/Product: osteocalcin #status experimental <MAT>  
F:60/Modified site: 4-hydroxyproline (Pro) #status absent  
F:68/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:72/75/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:74-80/Disulfide bonds: #status experimental

Query Match 51.0%; Score 25; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred.No. 3.7e-20;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
Db 76 LNPDCDELADHIGFOEAYRRFGPV 100

RESULT 4  
GEBQ

osteocalcin precursor - bovine  
N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing protein  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 24-Apr-1984 #sequence\_revision 07-Apr-1994 #text\_change 22-Jun-1999  
C:Accession: S12653; A03303; S08693  
R:Kiefer, M.C.; Saphire, A.C.S.; Bauer, D.M.; Barr, P.J.  
Nucleic Acids Res. 18, 1909, 1990  
A:title: The cDNA and derived amino acid sequences of human and bovine bone Gla protein.

A:Reference number: S12652; M01D:90245603; PMID:2336375  
A:Accession: S12653  
A:Molecule type: mRNA  
A:Residues: 1-100 <RTE>

A:Cross-references: EMBL:X53699; NID:9719; PIDN:CAA37737.1; PID:9720  
A:Note: alternative splicing may produce a sequence lacking residues 33-34  
R:Price, P.A.; Poser, J.W.; Raman, N.  
Proc. Natl. Acad. Sci. U.S.A. 73, 3374-3375, 1976  
A:title: Primary structure of the gamma-carboxyglutamic acid-containing protein from  
A:Reference number: A03303; M01D:77036749; PMID:1068450  
A:Accession: A03303  
A:Molecule type: protein  
A:Residues: 52-100 <PRI>  
C:Comment: This protein, isolated from bone, binds strongly to apatite.  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
F:1-51/Domain: signal sequence #status predicted <SIG>  
F:52-100/Product: osteocalcin #status experimental <MAT>  
F:60/Modified site: 4-hydroxyproline (Pro) #status experimental  
F:68/72/75/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:74-80/Disulfide bonds: #status experimental

Query Match 51.0%; Score 25; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred.No. 3.7e-20;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
Db 76 LNPDCDELADHIGFOEAYRRFGPV 100

RESULT 5  
GECT

osteocalcin - cat  
N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing prote  
C:Species: Felis silvestris catus (domestic cat)  
C>Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 06-Sep-1996  
C:Accession: A03304  
R:Shimomura, H.; Kanai, Y.; Sanada, K.  
J. Biochem. 96, 405-411, 1984  
A:title: Primary structure of cat osteocalcin.  
A:Reference number: A03304; M01D:85054706; PMID:6334077  
A:Accession: A03304  
A:Molecule type: protein  
A:Residues: 1-49 <SHI>  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
F:1-49/Domain: signal sequence #status predicted <SIG>  
F:50/Modified site: 4-hydroxyproline (Pro) #status experimental  
F:17,21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:23-29/Disulfide bonds: #status experimental

Query Match 30.6%; Score 15; DB 1; Length 49;  
Best Local Similarity 100.0%; Pred.No. 1.6e-09;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFO 39  
Db 25 LNPDCDELADHIGFO 39

RESULT 6  
GECB

osteocalcin precursor - chicken  
N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing prote  
C:Species: Gallus gallus (chicken)  
C>Date: 15-Oct-1982 #sequence\_revision 15-Aug-1997 #text\_change 22-Jun-1999  
C:Accession: I50700; A03305  
R:Neugebauer, B.M.; Moore, M.A.; Broess, M.; Gerstenfeld, L.C.; Hauschka, P.V.  
J. Bone Miner. Res. 10, 157-163, 1995  
A:title: Characterization of structural sequences in the chicken osteocalcin gene: ex  
A:Reference number: I50700; M01D:95266465; PMID:7747623  
A:Accession: I50700  
A:Status: preliminary; translated from GB/EMBL/DBJ



A:Molecule type: mRNA  
A:Residues: 1-97 <NEU>  
A:Cross-references: EMBL:U10578; NID:g1008455; PIDN:AA78809.1; PID:g5595408  
R:Cart, S.A.; Hauschka, P.V.; Bismann, K.  
J. Biol. Chem. 256, 9944-9950, 1981  
A:Title: Gas chromatographic mass spectrometric sequence determination of osteocalcin, a  
A:Reference number: A03305; MUID:82007831; PMID:6792200  
A:Accession: A03305  
A:Molecule type: protein  
A:Residues: 49-63, 'I', '65-77, 'N', '79-82, 'E', '85-90, 'Q', '91-97 <CAR>  
C:Comment: The gamma-carboxyglutamic acid residues formed by vitamin K-dependent posttra  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid  
F:49-97/Product: osteocalcin #status experimental <MAT>  
F:65,69,72/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:71-77/Disulfide bonds: #status predicted

Query Match 28.6%; Score 14; DB 1; Length 97;  
Best Local Similarity 100.0%; Pred. No. 3.7e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IGFQEAAYRRYGPV 49  
DB 84 IGFQEAAYRRYGPV 97  
|||||

RESULT 7  
GERT  
N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing protein  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 22-Jun-1999  
C:Accession: A31856; A31419; A32324; A25167; A25471  
R:Toom, K.; Rutledge, S.J.C.; Buenaqa, R.F.; Rodan, G.A.  
Biochemistry 27, 8521-8526, 1988  
A:Title: Characterization of the rat osteocalcin gene: stimulation of promoter activity  
A:Reference number: A31856; MUID:89118266; PMID:3265336  
A:Accession: A31856  
A:Molecule type: DNA  
A:Residues: 1-99 <TOO>  
A:Cross-references: GB:M23637; NID:g340986; PIDN:AAA1761.1; PID:g514962  
R:Theofan, G.; Haberstroh, L.M.; Price, P.A.  
DNA 8, 213-221, 1989  
A:Title: Molecular structure of the rat bone Gla protein gene and identification of puta  
A:Reference number: A31419; MUID:89251082; PMID:2785507  
A:Accession: A31419  
A:Molecule type: DNA  
A:Residues: 1-99 <THE>  
A:Cross-references: GB:M25490; NID:g576530; PIDN:AAA53280.1; PID:g576531  
R:lian, J.; Stewart, C.; Puchacz, E.; Mackowiak, S.; Shalhoub, V.; Collart, D.; Zambetti  
Proc. Natl. Acad. Sci. U.S.A. 86, 1143-1147, 1989  
A:Title: Structure of the rat osteocalcin gene and regulation of vitamin D-dependent exp  
A:Reference number: A32324; MUID:89145200; PMID:2784002  
A:Accession: A32324  
A:Molecule type: DNA  
A:Residues: 1-99 <LIA>  
A:Cross-references: GB:J04500; NID:g205863; PIDN:AAA1764.1; PID:g205864  
R:Pan, L.C.; Price, P.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6109-6113, 1985  
A:Title: The propeptide of rat bone gamma-carboxyglutamic acid protein shares homology w  
A:Reference number: A25167; MUID:85298305; PMID:3875856  
A:Accession: A25167  
A:Molecule type: mRNA  
A:Residues: 1-99 <PAN>  
A:Cross-references: GB:M11777; NID:g203147; PIDN:AAA0816.1; PID:g203148  
R:Celstey, A.J.; Rosen, V.; Buecker, J.L.; Kriz, R.; Wang, E.A.; Wozney, J.M.  
EMBO J. 5, 1885-1890, 1986  
A:Title: Isolation of the human gene for bone gla protein utilizing mouse and rat cDNA c  
A:Reference number: A91045; MUID:87004555; PMID:3019668  
A:Accession: A25471  
A:Molecule type: mRNA  
A:Residues: 1-99 <CEL>  
A:Cross-references: GB:X04141; NID:g55826; PIDN:CMA27761.1; PID:g55827

C:Genetics:  
A:Introns: 22/1; 33/1; 56/2  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
F:1-21/Domain: signal sequence #status predicted <SIS>  
F:22-49/Domain: propeptide #status predicted <PRO>  
F:50-59/Product: osteocalcin #status predicted <CON>  
F:58/Modified site: 4-hydroxyproline (Pro) #status predicted  
F:66,70,73/Modified site: gamma-carboxyglutamic acid (Glu) #status predicted  
F:72-78/Disulfide bonds: #status predicted

Query Match 22.4%; Score 11; DB 1; Length 99;  
Best Local Similarity 100.0%; Pred. No. 7.1e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELDHICFQ 39  
DB 78 CDELDHICFQ 88  
|||||

RESULT 8  
A61280  
N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing prote  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 12-May-1994 #sequence\_revision 02-Jun-1994 #text\_change 06-Sep-1996  
C:Accession: A61280  
R:Virdi, A.S.; Willis, A.C.; Hauschka, P.V.; Triffitt, J.T.  
Biochem. Soc. Trans. 19, 373S, 1991  
A:Title: Primary aminoacid sequence of rabbit osteocalcin.  
A:Reference number: A61280; MUID:92175242; PMID:1794506  
A:Accession: A61280  
A:Molecule type: protein  
A:Residues: 1-49 <VIR>  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline; pyroglutamic  
F:1/Modified site: pyroglutamate carboxylic acid (Glu) #status experimental  
F:9/Modified site: 4-hydroxyproline (Pro) #status experimental  
F:17,21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:33-39/Disulfide bonds: #status predicted

Query Match 20.4%; Score 10; DB 1; Length 49;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELAD 34  
DB 25 LNPDCDELAD 34  
|||||

RESULT 9  
T42606  
probable transcription activator - equine herpesvirus 4 (strain NS80567)  
C:Species: equine herpesvirus 4  
A:Variety: strain NS80567  
C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T42606  
R:Reilford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.  
J. Gen. Virol. 79, 1197-1203, 1998  
A:Title: The DNA sequence of equine herpesvirus-4.  
A:Reference number: Z22173; MUID:98264497; PMID:9603335  
A:Accession: T42606  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-536 <TEL>  
A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59582.1; PID:g2606010  
A:Experimental source: strain NS80567  
C:Genetics:  
A:Note: 63  
C:Superfamily: varicella-zoster virus gene 63 protein; RING finger homology  
F:5-53/Domain: RING finger homology <RRN>

Query Match 16.3%; Score 8; DB 2; Length 536;

Best Local Similarity 100.0%; Pred. No. 0.62; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 9 PVPDPPL 16  
Db 486 PVPDPPL 493

RESULT 10  
E97419  
hypothetical protein AGR\_C\_880 [imported] - Agrobacterium tumefaciens (strain C58, Cerec  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: E97419  
R:Goodner, B.; Hinkley, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; PMID:11743194  
A:Accession: E97419  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-275 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK6310.1; PID:q15155426; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C\_880  
A:Map position: circular chromosome

Query Match 14.3%; Score 7; DB 2; Length 275;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 38 FOEAYR 44  
Db 167 FOEAYR 173

RESULT 11  
B84226  
3-hydroxyacyl-CoA dehydrogenase [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: B84226  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: B84226  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-286 <STO>  
A:Cross-references: GB:AE004437; NID:q10580268; PIDN:AAG19174.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: hbd1  
C:Superfamily: 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenase homology

Query Match 14.3%; Score 7; DB 2; Length 286;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 31 ELADHG 37  
Db 234 ELADHG 240

RESULT 12  
G83544  
probable transcription regulator PA0815 [imported] - Pseudomonas aeruginosa (strain PA01  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: G83544  
R:Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: G83544  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-317 <STO>  
A:Cross-references: GB:AE004516; GB:AE004091; NID:q9946699; PIDN:AAG04204.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA0815

Query Match 14.3%; Score 7; DB 2; Length 317;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 QWLGAPV 10  
Db 42 QWLGAPV 48

RESULT 13  
I38975  
nuclear orphan receptor LXR-alpha - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Feb-1996 #sequence\_revision 16-Feb-1996 #text\_change 20-Sep-1999  
C:Accession: I38975  
R:Willly, P.J.; Umesono, K.; Ong, E.S.; Evans, R.M.; Heyman, R.A.; Mangelsdorf, D.J.  
Genes Dev. 9, 1033-1045, 1995  
A:Title: LXR, a nuclear receptor that defines a distinct retinoid response pathway.  
A:Reference number: I38975; MUID:95262897; PMID:7744246  
A:Accession: I38975  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-447 <RES>  
A:Cross-references: EMBL:U22662; NID:q726512; PIDN:AA85856.1; PID:q726513  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: zinc finger  
E:96-367/Domain: erba transforming protein homology <ERBA>

Query Match 14.3%; Score 7; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 WLGAPV 11  
Db 4 WLGAPV 10

RESULT 14  
WMBE6E  
65K early nonstructural protein - human cytomegalovirus (strain AD169)  
N:Alternate names: UL84 protein  
C:Species: human cytomegalovirus, human herpesvirus 5  
A>Note: host Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 16-Jul-1999  
C:Accession: S09848  
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T  
M.; Barrett, B.G.  
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovi  
A:Reference number: S09749; MUID:90269039; PMID:2161319  
A:Accession: S09848  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-586 <CHR>  
A:Cross-references: EMBL:X17403; NID:q59591; PIDN:CA33358.1; PID:q1780862  
A>Note: this sequence was submitted to the EMBL Data Library, December 1989  
C:Superfamily: human cytomegalovirus 65K early nonstructural protein

C:Keywords: leucine zipper; nonstructural protein  
F:114-135/Region: leucine zipper motif  
F:324-372/Region: leucine zipper motif

Query Match 14.3%; Score 7; DB 1; Length 586;  
Best Local Similarity 100.0%; Pred. No. 8.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFGP 48  
DB 443 YRRFGP 449

## RESULT 15

MMBERE

65K early nonstructural protein - human cytomegalovirus (strain Towne)  
C:Species: human cytomegalovirus, human herpesvirus 5  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 16-Jul-1999

C:Accession: A41808  
R:He, Y.S.; Xu, L.; Huang, E.S.

J. Virol. 66, 1096-1108, 1992

A:Title: Characterization of human cytomegalovirus UL84 early gene and identification of  
A:Reference number: A41808; MUID:92114132; PMID:1309892

A:Accession: A41808

A:Molecule type: mRNA

A:Residues: 1-587 <HEX>

A:Cross-references: GB:M81432; NID:9330541; PIDN:AAA45947.1; PID:9330542

C:Superfamily: human cytomegalovirus 65K early nonstructural protein

C:Keywords: leucine zipper; nonstructural protein

F:114-135/Region: leucine zipper motif  
F:325-373/Region: leucine zipper motif

Query Match 14.3%; Score 7; DB 1; Length 587;  
Best Local Similarity 100.0%; Pred. No. 8.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFGP 48  
DB 444 YRRFGP 450

## RESULT 16

TJ1345

hypothetical protein G01D9.5 - Caenorhabditis briggsae

C:Species: Caenorhabditis briggsae

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 01-Dec-2000

C:Accession: TJ1345

R:Waterston, R.

submitted to the EMBL Data Library, April 1996

A:Description: The C. briggsae genome sequencing project.

A:Reference number: Z21010

A:Accession: TJ1345

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4767 <WAT>

A:Cross-references: EMBL:U56248; NID:91293789; PID:91293790; PIDN:AAA98699.1

C:Genetics:

A:Insertions: 269/3; 341/3; 853/1; 920/2; 4452/3; 4534/3; 4592/3; 4654/3; 4670/3; 4707/2; 4

A:Note: G01D9.5

C:Superfamily: acyl carrier protein homology; acetate-CoA ligase homology

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:3472-3541/Domain: acyl carrier protein homology <ACPT>

F:4039-4427/Domain: acetate-CoA ligase homology <ACLI>

F:4447-4514/Domain: acyl carrier protein homology <ACPT>

F:2210/3505/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 14.3%; Score 7; DB 2; Length 4767;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 DELADHI 36  
|||||||

DB 35 DELADHI 41

## RESULT 17

C71923

ribosomal protein S21 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 24-Sep-1999

C:Accession: C71923

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, O.; Taylor, D.E.; Voris, G.F.

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: C71923

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-70 <ARN>

A:Cross-references: GB:AE001484; GB:AE001439; NID:94155043; PIDN:AAD06085.1; PID:9415

A:Experimental source: strain J99

C:Genetics:

A:Gene: rpsu

C:Superfamily: Escherichia coli ribosomal protein S21

Query Match 12.2%; Score 6; DB 2; Length 70;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EAYRRF 45  
DB 14 EAYRRF 19

## RESULT 18

B64590

ribosomal protein S21 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 13-Aug-1999

C:Accession: B64590

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpi, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: B64590

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-70 <TON>

A:Cross-references: GB:AE000570; GB:AE000511; NID:92313672; PIDN:AAD07628.1; PID:9231

C:Superfamily: Escherichia coli ribosomal protein S21

Query Match 12.2%; Score 6; DB 2; Length 70;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EAYRRF 45  
DB 14 EAYRRF 19

## RESULT 19

S62333

L71-1 protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 19-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 24-Sep-1999

C:Accession: S62333; S62342

R:Wright, L.G.; Chen, T.; Thummel, C.S.; Guild, G.M.

J. Mol. Biol. 255, 387-400, 1996

A:Title: Molecular characterization of the 71E late puff in Drosophila melanogaster r

A:Reference number: S62333; MUID:96152797; PMID:8568884

A:Accession: S62333  
 A:Molecule type: DNA  
 A:Residues: 1-100 <WR>  
 A:Cross-references: EMBL:U23836; NID:g939996; PIDN:AAA74176.1; PID:g939997  
 A:Accession: S62342  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-100 <WRW>  
 A:Cross-references: EMBL:U24095; NID:g775224; PIDN:AAA65109.1; PID:g775225  
 C:Genetics:  
 A:Gene: l7l-1  
 A:Cross-references: FlyBase:FBgn0004588  
 A:Introns: 12/1; 79/1  
 C:Superfamily: L71-10 protein  
 Query Match 12.2%; Score 6; DB 2; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 28 DCDDEL 33  
 DB 24 DCDDEL 29  
 RESULT 20  
 E86564  
 Yidd family [imported] - Chlamydia pneumoniae (strain J138)  
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
 C:Accession: E86564  
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:20330349; PMID:10871362  
 A:Accession: E86564  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-103 <STO>  
 A:Cross-references: GB:BA000008; NID:g8978964; PIDN:BAA98799.1; GSPDB:GN00142  
 C:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: yidd  
 Query Match 12.2%; Score 6; DB 2; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YLXQWL 6  
 DB 20 YLXQWL 25  
 RESULT 21  
 H72060  
 Yidd family - Chlamydia pneumoniae (strain CWL029)  
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000  
 C:Accession: H72060  
 R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
 Nature Genet. 21, 385-389, 1999  
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
 A:Reference number: A72000; MUID:99206606; PMID:10192388  
 A:Accession: H72060  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-103 <ARN>  
 A:Cross-references: GB:AE001643; GB:AE001363; NID:g4376876; PIDN:AAD18731.1; PID:g437688  
 A:Experimental source: strain CWL029  
 C:Genetics:  
 A:Gene: yidd  
 Query Match 12.2%; Score 6; DB 2; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 21;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YLXQWL 6  
 DB 20 YLXQWL 25  
 RESULT 22  
 E81606  
 conserved hypothetical protein CP0156 [imported] - Chlamydia pneumoniae (strain A  
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: E81606  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
 , C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39  
 A:Reference number: A81500; MUID:20150255; PMID:10684935  
 A:Accession: E81606  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-106 <REA>  
 A:Cross-references: GB:AE002177; GB:AE002161; NID:g7189090; PIDN:AAF38037.1; PID:g718  
 C:Genetics:  
 A:Gene: CP0156  
 Query Match 12.2%; Score 6; DB 2; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YLXQWL 6  
 DB 23 YLXQWL 28  
 RESULT 23  
 S38261  
 myohemerythrin - polychaete (Nereis diversicolor)  
 C:Species: Nereis diversicolor (sandworm)  
 C:Date: 19-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 22-Oct-1999  
 C:Accession: S38261  
 R:Demyanck, S.; Li, K.W.; van der Schors, R.; Dhainaut-Courtois, N.  
 Eur. J. Biochem. 217, 151-156, 1993  
 A:Title: Amino acid sequence of the small cadmium-binding protein (MP II) from Nereis  
 A:Reference number: S38261; MUID:94039032; PMID:8223553  
 A:Accession: S38261  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-119 <DEM>  
 C:Superfamily: hemerythrin  
 C:Keywords: Iron; Oxygen carrier  
 F;25,54,58,74,78,107,112/Binding site: 2Fe-O cluster (His, His, Glu, His, His, His, A  
 Query Match 12.2%; Score 6; DB 2; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 LGAPVP 11  
 DB 88 LGAPVP 93  
 RESULT 24  
 T18052  
 DEAH box protein a550R - Chlorella virus PBCV-1  
 C:Species: Chlorella virus PBCV-1  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T18052  
 R:Graves, M.V.; Van Etten, J.L.  
 submitted to the EMBL Data Library, May 1999  
 A:Reference number: Z18806  
 A:Accession: T18052

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-136 <GRA>  
 A:Cross-references: EMBL:U42580; NID:94028896; PIDN:AAC9697.1  
 A:Experimental source: specific host *Chlorella* strain NC64A  
 C:Genetics:  
 A:Note: a550R  
 C:Superfamily: *Chlorella* virus PBCV-1 DEAH box protein a550R

Query Match 12.2%; Score 6; DB 2; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 LGAPVP 11  
 |||||  
 Db 12 LGAPVP 17

RESULT 25  
 UDCH  
 cystatin precursor - chicken  
 N:Alternate names: cystatin I; cysteine proteinase inhibitor; egg-white cystatin  
 C:Species: *Gallus gallus* (chicken)  
 C>Date: 03-Aug-1984 #sequence\_revision 12-Apr-1996 #text\_change 29-Oct-1999  
 C:Accession: A34456; A01274; S48159; S04008; JN0789  
 R:Coella, R.; Sakaguchi, Y.; Nagase, H.; Bird, J.W.C.  
 J. Biol. Chem. 264, 17164-17169, 1989  
 A:Title: Chicken egg white cystatin. Molecular cloning, nucleotide sequence, and tissue  
 A:Reference number: A34456; MUID:90008873; PMID:2793849  
 A:Accession: A34456  
 A:Molecule type: mRNA  
 A:Residues: 1-139 <COL>  
 A:Cross-references: GB:J05077; NID:g211714; PIDN:AAA0744.1; PID:g211715  
 R:Schwabe, C.; Anastasi, A.; Crow, H.; McDonald, J.K.; Barrett, A.J.  
 Biochem. J. 217, 813-817, 1984  
 A:Title: Cystatin. Amino acid sequence and possible secondary structure.  
 A:Reference number: A01274; MUID:84178305; PMID:6712597  
 A:Accession: A01274  
 A:Molecule type: protein  
 A:Residues: 24-139 <SCH>  
 R:Turk, V.; Brzin, J.; Longier, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983  
 A:Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystat  
 A:Reference number: S01461; MUID:84110059; PMID:6662498  
 A:Accession: S01461  
 A:Molecule type: protein  
 A:Residues: 24-139 <TUR>  
 R:Anastasi, A.; Brown, M.A.; Kembhavi, A.A.; Nicklin, M.J.H.; Sayers, C.A.; Sunter, D.C.  
 Biochem. J. 211, 129-138, 1983  
 A:Title: Cystatin, a protein inhibitor of cysteine proteinases. Improved purification fr  
 A:Reference number: A37514; MUID:83256421; PMID:6409085  
 A:Contents: annotation; characterization of protein  
 R:Gund, A.; Loiberg, H.; Barrett, A.J.  
 FEBS Lett. 170, 370-374, 1984  
 A:Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
 A:Reference number: S01462  
 A:Contents: annotation; disulfide bonds  
 R:Auerwald, E.A.; Neigler, D.K.; Schilze, A.J.; Engh, R.A.; Genenger, G.; Machleidt, W.  
 Eur. J. Biochem. 224, 407-415, 1994  
 A:Title: Production, inhibitory activity, folding and conformational analysis of an N-te  
 A:Reference number: S48159; MUID:95010016; PMID:7925354  
 A:Accession: S48159  
 A:Status: preliminary  
 A:Molecule type: protein  
 C:Species: *Rattus norvegicus* (Norway rat)  
 R:Lieber, B.; Krieglstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.  
 FEBS Lett. 248, 162-168, 1989  
 A:Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.  
 A:Reference number: S04008; MUID:89252033; PMID:2721673  
 A:Accession: S04008  
 A:Molecule type: protein  
 A:Residues: 97-114 <LAB>  
 R:Coella, R.; Bird, J.W.C.

Gene 130, 175-181, 1993  
 A:Title: Isolation and characterization of the chicken cystatin-encoding gene. Mappin  
 A:Reference number: JN0789; MUID:93366172; PMID:8359684  
 A:Accession: JN0789  
 A:Molecule type: DNA  
 A:Residues: 1-139 <CO2>  
 A:Cross-references: GB:M95725  
 A:Note: authors failed to translate the codon for residue 115-Tyr  
 C:Comment: This protein binds tightly to and inhibits a variety of cysteine proteinas  
 C:Genetics:  
 A:Gene: Csn  
 A:Introns: 76/3; 114/3  
 C:Superfamily: cystatin; cystatin homology  
 C:Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein  
 F:1-23/Domain: signal sequence #status predicted <IG>  
 F:24-139/Product: cystatin, long form #status experimental <CYLF>  
 F:30-139/Domain: cystatin homology <CYS>  
 F:32-139/Product: cystatin, short form #status experimental <CVSFR>  
 F:76-80/Region: inhibitory #status predicted  
 F:94-104, 118-138/Disulfide bonds: #status experimental  
 F:103/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 12.2%; Score 6; DB 1; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 LGAPVP 11  
 |||||  
 Db 31 LGAPVP 36

RESULT 26  
 T12144  
 hypothetical protein 143 - fava bean mitochondrion  
 C:Species: mitochondrion *Vicia faba* (fava bean)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Sep-1999  
 C:Accession: T12144  
 R:Sheepers, D.G.; Luo, H.; Boult, M.  
 Plant Sci. 129, 203-212, 1997  
 A:Title: Variant mitochondrial transcripts of a broad bean line are associated with t  
 A:Reference number: Z17435  
 A:Accession: T12144  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-143 <SCH>  
 A:Cross-references: EMBL:L36945; NID:9558142; PID:9558143  
 C:Genetics:  
 A:Genome: mitochondrion  
 C:Superfamily: fava bean mitochondrion hypothetical protein 143  
 C:Keywords: mitochondrion

Query Match 12.2%; Score 6; DB 2; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 APVPYP 13  
 |||||  
 Db 55 APVPYP 60

RESULT 27  
 PC4154  
 Zn-alpha 2-glycoprotein - rat (fragment)  
 C:Species: *Rattus norvegicus* (Norway rat)  
 C>Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 11-Jan-2000  
 C:Accession: PC4154  
 R:Uita, J.A.; Fuenyo, A.; Balbin, M.; Velasco, G.; Pendas, A.M.; Lopez-Otin, C.  
 Gene 169, 233-236, 1996  
 A:Title: Alternative splicing gives rise to two novel long isoforms of Zn-alpha2-glyc  
 A:Reference number: PC4154; MUID:96194808; PMID:86477433  
 A:Accession: PC4154  
 A:Molecule type: mRNA  
 A:Residues: 1-164 <URI>

A:Cross-references: EMBL:X86178  
 A:Experimental source: liver  
 C:Comment: It is involved in the transport of nonpolymorphic substances or intercellular  
 D:  
 C:Genetics:  
 A:Gene: Zn-alpha 2-gp  
 A:introns: 7/1; 77/2; 83/1  
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
 C:Keywords: glycoprotein; immune response; immunoglobulin; liver  
 F:79/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.2%; Score 6; DB 2; Length 164;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 PYPDPL 16  
 Db 48 PYPDPL 53

RESULT 28  
 S03744  
 hypothetical protein - Calothrix sp.  
 C:Species: Calothrix sp.  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 18-Jun-1993  
 C:Accession: S03744  
 R:Parsoot, C.; Mazel, D.  
 Mol. Microbiol. 1, 45-52, 1987  
 A:Title: Cloning and nucleotide sequence of the thrB gene from the cyanobacterium Caloth  
 A:Reference number: S03743; MUID:88260883; PMID:2838727  
 A:Accession: S03744  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-169 <PAR>

Query Match 12.2%; Score 6; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EAYRRF 45  
 Db 43 EAYRRF 48

RESULT 29  
 S44737  
 CO2C2.6 protein - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Sep-1997  
 C:Accession: S44737  
 R:Wilson, R.  
 submitted to the EMBL Data Library, September 1993  
 A:Description: Sequence of the C. elegans cosmid C02c2.  
 A:Reference number: S44737  
 A:Accession: S44737  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-171 <WIL>  
 A:Cross-references: EMBL:L23649; NID:g388566; PID:g388567  
 C:Genetics:  
 A:introns: 79/3; 118/3

Query Match 12.2%; Score 6; DB 2; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 DELADH 35  
 Db 109 DELADH 114

2 RESULT 30  
 AH1075

probable fimbrial chain sthd [imported] - Salmonella enterica subsp. enterica serovar  
 C:Species: Salmonella enterica subsp. enterica serovar Typh  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 01-Mar-2002  
 C:Accession: AH1075  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Church  
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
 S.; Mole, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
 A:Reference number: AB0502; PMID:11577608  
 A:Accession: AH1075  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-185 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD03423.1; PID:g16505690; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: sthd  
 C:Superfamily: conserved hypothetical protein b1503

Query Match 12.2%; Score 6; DB 2; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 VPYPDP 15  
 Db 99 VPYPDP 104

RESULT 31  
 AB2392  
 hypothetical protein all4690 [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AB2392  
 R:kaneko, T.; Nakamura, Y.; Molk, C.P.; Kurita, T.; Sasamoto, S.; Watanabe, A.; Itigu  
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AB2392  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-214 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA076389.1; PID:g17133827; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all4690

Query Match 12.2%; Score 6; DB 2; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 QWLGAP 9  
 Db 196 QWLGAP 201

RESULT 32  
 D81201  
 cell division protein FtsQ NMB0425 [imported] - Neisseria meningitidis (strain MC58 s  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C:Accession: D81201  
 R:Petelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.  
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masiugnani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Frazer, C.M.; Moxon, E.R.; Rappuoli, R.;  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: AB1000; MUID:20175755; PMID:10710307

A:Accession: D81201  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-235 <TEXT>  
A:Cross-references: GB:AE002398; GB:AE002098; NID:g7225640; PIDN:AAF40863.1; PID:g722564  
A:Experimental source: serogroup B, strain MCS8  
C:Genetics:  
A:Gene: NMB0425  
C:Superfamily: cell division protein ftsQ

Query Match 12.2%; Score 6; DB 2; Length 235;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 QEAYRR 44  
|||||  
DB 70 QEAYRR 75

## RESULT 33

G81776  
cell division protein NMA2059 [imported] - Neisseria meningitidis (strain Z2491 serogrou  
C:Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: G81776  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
R.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Mature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A:Reference number: A81775; MUID:20222536; PMID:10761919  
A:Accession: G81776  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-242 <PAR>  
A:Cross-references: GB:AL162758; GB:AL157959; NID:g7380672; PIDN:CAB85277.1; PID:g738068  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: ftsQ; NMA2059  
C:Superfamily: cell division protein ftsQ

Query Match 12.2%; Score 6; DB 2; Length 242;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 QEAYRR 44  
|||||  
DB 77 QEAYRR 82

## RESULT 34

AH3267  
acetyltransferase (EC 2.3.1.-) [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AH3267  
R:DeVecchio, V.G.; Kaparatil, V.; Redkar, R.J.; Patra, G.; Muir, C.; Los, T.; Ivanova,  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AH3267  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-252 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL51307.1; PID:g17982002; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0125  
A:Map position: I  
C:Keywords: acyltransferase

Query Match 12.2%; Score 6; DB 2; Length 252;  
Best Local Similarity 100.0%; Pred. No. 48;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 37 GFOEAY 42  
|||||  
DB 237 GFOEAY 242

## RESULT 35

T35168  
probable transcription regulator - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jan-2000  
C:Accession: T35168  
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, July 1998  
A:Reference number: Z21570  
A:Accession: T35168  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-264 <SEE>  
A:Cross-references: EMBL:AL031107; PIDN:CAA19948.1; GSPDB:GN00070; SCOEDB:SC5A7.19c  
C:Genetics:  
A:Gene: SCOEDB:SC5A7.19c  
C:Superfamily: acetate operon repressor

Query Match 12.2%; Score 6; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 FOEAYR 43  
|||||  
DB 153 FOEAYR 158

## RESULT 36

A87646  
hypothetical protein CC3203 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: A87646  
R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: A87646  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-289 <STO>  
A:Cross-references: GB:AE005673; NID:g13424881; PIDN:AAK25165.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3203

Query Match 12.2%; Score 6; DB 2; Length 289;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 QEAYRR 44  
|||||  
DB 214 QEAYRR 219

## RESULT 37

C69111  
guinolinate synthetase - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Jul-1999  
C:Accession: C69111  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T  
R.; Qiu, D.; Spadefora, R.; Vicalire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997  
 A>Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: C69111  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-304 <MT>  
 A:Cross-references: GB:A6000936; GB:A6000666; NID:g2622959; PIDN:AA868293.1; PID:g262296  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH1827  
 C:Superfamily: quinolinate synthetase A

Query Match 12.2%; Score 6; DB 2; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADH1 36  
 |||||  
 Db 208 ELADH1 213

RESULT 38  
 F75473  
 conserved hypothetical protein - *Deinococcus radiodurans* (strain R1)  
 C:Species: *Deinococcus radiodurans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 A:Accession: F75473  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A>Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: F75473  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-332 <WH1>  
 A:Cross-references: GB:A6001935; GB:A6000513; NID:96458517; PIDN:AAF10387.1; PID:9645852  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR0809  
 A:Map position: 1

Query Match 12.2%; Score 6; DB 2; Length 332;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPP 11  
 |||||  
 Db 204 LGAPP 209

RESULT 39  
 E69964  
 NADH-dependent flavin oxidoreductase homolog yqjM - *Bacillus subtilis*  
 C:Species: *Bacillus subtilis*  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 A:Accession: E69964  
 R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beter  
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Gall  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinots,  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
 Rieger, M.; Rivolta, C.; Rooha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 akouchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: E69964  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-338 <K0>  
 A:Cross-references: GB:299116; GB:AL009126; NID:g2634723; PIDN:CAB14314.1; PID:g26348  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yqjM  
 C:Superfamily: NADPH dehydrogenase chain OYE2

Query Match 12.2%; Score 6; DB 2; Length 338;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ADHGP 38  
 |||||  
 Db 229 ADHGP 234

RESULT 40  
 C70532  
 hypothetical protein Rv2712c - *Mycobacterium tuberculosis* (strain H37RV)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 A:Accession: C70532  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
 Connor, R.A.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd,  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.  
 A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: C70532  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-352 <COL>  
 A:Cross-references: GB:296072; GB:AL123456; NID:g3261793; PIDN:CAB09472.1; PID:e32110  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: Rv2712c

Query Match 12.2%; Score 6; DB 2; Length 352;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELAD 34  
 |||||  
 Db 51 CDELAD 56

RESULT 41  
 E64399  
 hypothetical protein MJ0797 - *Methanococcus jannaschii*  
 C:Species: *Methanococcus jannaschii*  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 A:Accession: E64399  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak  
 reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,  
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Moese  
 A>Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannasc*  
 A:Reference number: A64300; MUID:96337999; PMID:8688087  
 A:Accession: E64399  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-367 <BU>  
 A:Cross-references: GB:U67524; GB:U77117; NID:g2826321; PIDN:AA898792.1; PID:g1591494  
 C:Genetics:  
 A:Map position: FOR720661-721764



Query Match 12.2%; Score 6; DB 2; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYG 47  
 |||||  
 DB 178 YRRFYG 183

RESULT 42

S32173  
 Hypothetical protein 6 - *Myxococcus xanthus*  
 C:Species: *Myxococcus xanthus*  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: S32173  
 R:Botella, J.; Murillo, F.; Ruiz-vazquez, R.  
 Submitted to the EMBL Data Library, March 1993  
 A:Description: Nucleotide and deduced protein sequences of a carotenoid gene cluster in  
 A:Reference number: S32168  
 A:Accession: S32173  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-370 <R07>  
 A:Cross-references: EMBL:221955; NID:95775589; PIDN:CAA79960.1; PID:9288225

Query Match 12.2%; Score 6; DB 2; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 LADHIG 37  
 |||||  
 DB 202 LADHIG 207

RESULT 43

G70796  
 Hypothetical protein RV3726 - *Mycobacterium tuberculosis* (strain H37RV)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: G70796  
 R:Colet, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Nature 393, 537-544, 1998  
 A:Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A:Reference number: A70500; M0ID:98295987; PMID:9634230  
 A:Accession: G70796  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-397 <C0L>  
 A:Cross-references: GB:AL022121; GB:AL123456; NID:93261559; PIDN:CAA18048.1; PID:9296015  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: RV3726  
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology  
 F:21-373/Domain: long-chain alcohol dehydrogenase homology <LADH>

Query Match 12.2%; Score 6; DB 2; Length 397;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 CDELAD 34  
 |||||  
 DB 46 CDELAD 51

RESULT 44

A33396  
 beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - *mus musculus* (house mouse)  
 C:Species: *Mus musculus* (house mouse)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Jun-2000

C:Accession: A33396; A28799; JX0041; I49531; I49532  
 R:Holliis, G.F.; Douglas, J.G.; Shaper, N.L.; Shaper, J.H.; Stafford-Holliis, J.M.; Eva Biochem. Biophys. Res. Commun. 162, 1069-1075, 1989  
 A:Title: Genomic structure of murine beta-1,4-galactosyltransferase.  
 A:Reference number: A33396; M0ID:89350913; PMID:2504153

A:Accession: A33396  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-399 <H0L>

A:Cross-references: GB:M27922; NID:9341810; PIDN:AAA58744.1; PID:9609528  
 A:Note: the authors did not translate the codon GTT for residue 139  
 R:Shaper, N.L.; Holliis, G.F.; Douglas, J.G.; Kirsch, I.R.; Shaper, J.H.  
 J. Biol. Chem. 263, 10420-10428, 1988

A:Title: Characterization of the full length cDNA for murine beta-1,4-galactosyltrans  
 A:Reference number: A28799; M0ID:88273147; PMID:3134348

A:Accession: A28799  
 A:Molecule type: mRNA  
 A:Residues: 1-399 <SHA>  
 A:Cross-references: EMBL:J03880; NID:9192195; PIDN:AAA37297.1; PID:9309117  
 A:Note: It is uncertain whether Met-1 or Met-14 is the Initiator  
 R:Nakazawa, K.; Ando, T.; Kimura, T.; Narimatsu, H.  
 J. Biochem. 104, 165-168, 1988

A:Title: Cloning and sequencing of a full-length cDNA of mouse N-acetylglucosamine (b  
 A:Reference number: JX0041; M0ID:89033997; PMID:3141392

A:Accession: JX0041  
 A:Molecule type: mRNA  
 A:Residues: 1-399 <NAK>

A:Cross-references: GB:D00314; NID:9220340; PIDN:BA00216.1; PID:9220341  
 R:Shaper, J.H.; Holliis, G.F.; Shaper, N.L.  
 Biochimie 70, 1683-1688, 1988

A:Title: Evidence for two forms of murine beta-1,4-galactosyltransferase based on clo  
 A:Reference number: I49531; M0ID:89207607; PMID:3149531

A:Accession: I49531  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-63 <RES>

A:Cross-references: GB:M36289; NID:9192188; PIDN:AAA37294.1; PID:9192189  
 A:Accession: I49532  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 14-63 <RE2>

A:Cross-references: GB:M36289; NID:9192188; PIDN:AAA37295.1; PID:9192190  
 C:Genetics:  
 A:Introns: 139/1; 217/3; 280/2; 321/2; 356/2

C:Keywords: glycoprotein; glycosyltransferase; hexosyltransferase; transmembrane prot  
 F:25-44/Domain: transmembrane #status predicted <TM>  
 F:113/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.2%; Score 6; DB 2; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 IGFOEA 41  
 |||||  
 DB 231 IGFOEA 236

RESULT 45

T06000  
 aspartic proteinase homolog F17M5.250 - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 22-Oct-1999  
 C:Accession: T06000  
 R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hohenseil, J.; Mewes, H.W.; Mayer, K.  
 Submitted to the Protein Sequence Database, March 1999  
 A:Reference number: Z15263  
 A:Accession: T06000  
 A:Molecule type: DNA  
 A:Residues: 1-420 <BEV>  
 A:Cross-references: EMBL:AL035678; GSPDB:GN00062; ATSP:F17M5.250  
 A:Experimental source: cultivar Columbia; BAC clone F17M5  
 C:Genetics:  
 A:Gene: ATSP:F17M5.250  
 A:Map position: 4

A; Introns: 20/1; 36/2; 75/3; 150/2; 226/2; 273/3; 340/3; 382/1

Query Match 12.2%; Score 6; DB 2; Length 420;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 DCDELA 33  
|||||

Db 405 DCDELA 410

Search completed: December 4, 2002, 15:51:22  
Job time : 21 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: December 4, 2002, 15:48:06 ; Search time 11 Seconds  
(without alignments)  
184.758 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 49  
Sequence: 1 YLYQWLGAIPVPRDPLXPRR.....DELADHIGFOEAYRRYGPV 49

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	51.0	48	1 OSTC_DRONO	P15504 dromaius no
2	25	51.0	49	1 OSTC_HORSE	P83005 equus cabal
3	25	51.0	49	1 OSTC_MACFA	P02819 macaca fasc
4	25	51.0	49	1 OSTC_XENLA	P40147 xenopus lae
5	25	51.0	100	1 OSTC_BOVIN	P02820 bos taurus
6	25	51.0	100	1 OSTC_HUMAN	P02818 homo sapien
7	15	30.6	49	1 OSTC_FELCA	P02821 felis silve
8	14	28.6	49	1 OSTC_CANFA	P81455 canis fami
9	14	28.6	97	1 OSTC_CHICK	P02822 gallus gall
10	11	22.4	99	1 OSTC_CHICK	P04640 rattus norv
11	10	20.4	49	1 OSTC_RABIT	P39056 cryctolagus
12	7	14.3	447	1 NRH3_HUMAN	Q13133 homo sapien
13	7	14.3	586	1 ULB4_HCVNA	P16727 human cytom
14	7	14.3	587	1 ULB4_HCVNA	P29839 human cytom
15	7	14.3	1153	1 JAK1_BRARE	Q12990 brachydanto
16	6	12.2	70	1 RS21_HELPJ	Q92199 heliobacte
17	6	12.2	70	1 RS21_HELPJ	P56028 heliobacte
18	6	12.2	103	1 Y592_CHIPN	Q92742 chlamydia p
19	6	12.2	116	1 CYT_COTJA	P81061 coturnix co
20	6	12.2	119	1 MP2_NERDI	P80255 neretis dive
21	6	12.2	139	1 CYT_CHICK	P01038 gallus gall
22	6	12.2	171	1 YK66_CAEEL	P42474 caenorhabdi
23	6	12.2	304	1 NADA_METTH	C02785 methanobact
24	6	12.2	338	1 YQJM_BACSU	P54550 bacillus su
25	6	12.2	367	1 Y797_METJA	Q58207 methanococc
26	6	12.2	399	1 B4G1_MOUSE	P15535 m beta-1,4-
27	6	12.2	413	1 PURK_CORAN	Q44678 corynebacte
28	6	12.2	445	1 CLPX_ANASP	O8yqk7 arabidopsi
29	6	12.2	461	1 HMCS_ARATH	P54873 arabidopsi
30	6	12.2	471	1 SYE_YERPE	O82ck0 yerstinia pe
31	6	12.2	532	1 ICPO_HSEVB	P28990 equine hepr
32	6	12.2	553	1 NOEB_RHIME	O52893 rhizobium m
33	6	12.2	788	1 TRS1_HCVNA	P09695 human cytom

34	6	12.2	838	1 OS94_MOUSE	P48722 mus musculus
35	6	12.2	845	1 PPSA_AERPE	Q9yec5 aeropyrum p
36	6	12.2	934	1 MSH2_HUMAN	P43246 homo sapien
37	6	12.2	1246	1 SKIW_HUMAN	Q15477 homo sapien
38	6	12.2	1396	1 ITA2_DROME	P12080 drosophila
39	6	12.2	1912	1 PTPD_HUMAN	P23468 homo sapien
40	6	12.2	3172	1 ERY3_SACER	Q03133 saccharopol
41	5	10.2	69	1 YBA1_SCHPO	Q96vq1 schizosacch
42	5	10.2	76	1 Y066_HALNI	Q9hsu8 halobacter1
43	5	10.2	92	1 Y886_METTH	O26972 methanobact
44	5	10.2	105	1 NIGM_HUMAN	Q95178 homo sapien
45	5	10.2	106	1 Y086_CAEEL	Q09238 caenorhabdi

## ALIGNMENTS

RESULT 1	OSTC_DRONO	STANDARD:	PRT:	48 AA.
AC	P15504:			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-			
DE	protein) (BGP).			
GN	BGLAP.			
OS	Dromaius novaehollandiae (Emu).			
OC	Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Palaeognathae; Casuariiformes; Dromiidae;			
OC	Dromaius.			
OX	NCBI_Taxid-8790;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE-88134266; PubMed-3501719;			
RA	Hug N.L., Tseng A., Chapman G.E.;			
RT	"The amino acid sequence of Emu osteocalcin: gas phase sequencing of			
RT	Gla-containing proteins".			
RL	Biochem. Int. 15:271-277(1987).			
CC	- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS			
CC	STRONGLY TO APATITE AND CALCIUM.			
CC	- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K			
CC	DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE			
CC	BINDING OF CALCIUM.			
CC	- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN			
CC	FAMILY.			
DR	PIR: S02208; S02208.			
DR	InterPro: IPR002384; GLA_bone.			
DR	InterPro: IPR00294; VItK_dep_GLA.			
DR	Pfam: PF00594; gla; 1.			
DR	PRINTS: PR00002; GLABONE.			
DR	SMART: SM00069; GLA; 1.			
KW	PROSITE: PS00011; GLU_CARBOXYLATION; 1.			
FT	Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone.			
FT	MOD_RES 16 16			
FT	MOD_RES 20 20			
FT	MOD_RES 23 23			
FT	MOD_RES 28 28			
FT	DISULFID 22 22			
FT	SEQUENCE 48 AA; 5292 MW; 50A0DF3BFA8CFD CRC64;			
QY	25 LNPDCDELADHIGFOEAYRRYGPV 49			
Db	24 LNPDCDELADHIGFOEAYRRYGPV 48			
RESULT 2	OSTC_HORSE	STANDARD:	PRT:	49 AA.
ID	P83005:			
AC	P83005:			

```

DT 15-JUN-2002 (rel. 41, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
GN BGLAP.
OS Equus caballus (horse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE, HYDROXYLATION, AND GAMMA-CARBOXYGLUTAMIC ACIDS.
RC TISSUE=Bone;
RA Carstangjen B., Mattiez R., Amory H., Lepage O.M., Remy B.;
RT "Isolation and characterization of equine osteocalcin.";
RL Ann. Med. Vet. 0:0-0(2002).
CC -!- FUNCTION: Constitutes 1-2% of the total bone protein. It binds
CC strongly to apatite and calcium.
CC -!- PTM: Gamma-carboxyglutamic acid residues are formed by vitamin K
CC dependent carboxylation. These residues are essential for the
CC binding of calcium.
CC -!- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLUT_CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K;
KM Hydroxylation.
FT MOD_RES 9 9 HYDROXYLATION.
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 23 29 BY SIMILARITY.
SQ SEQUENCE 49 AA; 5732 MW; A5B82601AD12857F CRC64;

OY Query Match 51.0%; Score 25; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 1,6e-21;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 LNPDCELADHIGFOEAYRRFGPV 49
|||||
25 LNPDCELADHIGFOEAYRRFGPV 49

RESULT 3
OSTC_MACFA
ID OSTC_MACFA STANDARD: PRT; 49 AA.
AC P02819;
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
GN BGLAP.
OS Macaca fascicularis (Crab eating macaque) (Cynomolus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE.
RX MEDLINE=82182842; PubMed=6978733;
RA Hauschka P.V., Carr S.A., Blemann K.;
RT "Primary structure of monkey osteocalcin.";
RL Biochemistry 21:638-642(1982).
CC -!- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -!- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -!- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.

```

```

DR PIR: A03302; GEMKI.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLUT_CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;
KM Bone.
FT MOD_RES 9 9 HYDROXYLATION.
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 23 29 BY SIMILARITY.
SQ SEQUENCE 49 AA; 5743 MW; C20116014DOC4958 CRC64;

OY Query Match 51.0%; Score 25; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 1,6e-21;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 LNPDCELADHIGFOEAYRRFGPV 49
|||||
25 LNPDCELADHIGFOEAYRRFGPV 49

RESULT 4
OSTC_XENLA
ID OSTC_XENLA STANDARD: PRT; 49 AA.
AC P40147;
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE.
RX MEDLINE=96139691; PubMed=8567186;
RA Cancelli M., Williamson M.K., Price P.A.;
RT "Amino-acid sequence of bone Gla protein from the African clawed toad
RT Xenopus laevis and the fish Sparus aurata.";
RL Int. J. Pept. Protein Res. 46:419-423(1995).
CC -!- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -!- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -!- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLUT_CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone.
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 23 29 BY SIMILARITY.
SQ SEQUENCE 49 AA; 5360 MW; 7A9A6F6A12B6047 CRC64;

OY Query Match 51.0%; Score 25; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 1,6e-21;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 LNPDCELADHIGFOEAYRRFGPV 49
|||||
25 LNPDCELADHIGFOEAYRRFGPV 49

```

OSTC_BOVIN	5	RESULT
ID	OSTC_BOVIN	STANDARD: PRT: 100 AA.
AC	P02820.	
DT	21-JUL-1986 (Rel. 01, Created)	
DT	01-AUG-1990 (Rel. 15, Last sequence update)	
DT	15-JUN-2002 (Rel. 41, Last annotation update)	
DE	Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).	
GN	BGLAP.	
OS	Bos taurus (Bovine).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
OC	Bovidae; Bovinae; Bos.	
OX	NCBI_TaxID:9913;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	MEDLINE=90245603; PubMed=2336375;	
RA	Kiefer M.C., Saphire A.C.S., Bauer D.M., Barr P.J.;	
RT	"The cDNA and derived amino acid sequences of human and bovine bone Gla protein";	
RL	Nucleic Acids Res. 18:1909-1909(1990).	
RN	[2]	
RP	SEQUENCE OF 52-100.	
RX	MEDLINE=77036749; PubMed=1068450;	
RA	Price P.A., Poser J.W., Raman N.;	
RT	"Primary structure of the gamma-carboxyglutamic acid-containing protein from bovine bone.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 73:3374-3375(1976).	
CC	-1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS STRONGLY TO APATITE AND CALCIUM.	
CC	-1- PPM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE BINDING OF CALCIUM.	
CC	-1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN FAMILY.	
CC	-----	
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; X53699; CAA37737.1; -	
DR	EMBL; X51700; CAA35997.1; -	
DR	PIR; A03303; GEB0.	
DR	PIR; S12653; S12653.	
DR	InterPro: IPR002384; GLA Bone.	
DR	InterPro: IPR000294; VitK_dep_GLA.	
DR	Pfam; PF00594; gla; 1.	
DR	PRINTS; PRO0002; GLABONE.	
DR	SMART; SM0069; GLA; 1.	
DR	PROSITE; PS00011; GLU CARBOXYLATION; 1.	
KW	Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation; Bone; Signal.	
FT	SIGNAL	1 23
FT	PROPEP	24 51
FT	CHAIN	52 100
FT	MOD_RES	60 60
FT	MOD_RES	68 68
FT	MOD_RES	72 72
FT	MOD_RES	75 75
FT	DISULFID	74 80
SQ	SEQUENCE	100 AA; 11042 MW; 73015D1681B26219 CRC64;
Query Match	51.0%; Score 25; DB 1; Length 100;	
Best Local Similarity	100.0%; Pred. NO. 2; 9e-21;	
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

Oy	25	LNPDCDELADHIGFOEAYRRFGPV	49
Dd	76	LNPDCDELADHIGFOEATRRRTGVPV	100
RESULT 6			
ID	OSTC_HUMAN	STANDARD:	PRT: 100 AA.
AC	P02818.		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	01-AUG-1990 (Rel. 15, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Osteocalcin precursor (Gamma-carboxylglutamic acid-containing protein) (Bone GLA-protein) (BGP).		
Gn	BGLAP.		
Os	Homo sapiens (Human).		
Oc	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Ox	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
CX	NCHI_TaxID=9606;		
RN	[1]		
RX	SEQUENCE FROM N.A.		
RX	MEDLINE=90245603; PubMed=2336375;		
RA	Kiefer M.C., Saphire A.C.S., Bauer D.M., Barr P.J. ;		
RT	"The cDNA and derived amino acid sequences of human and bovine bone GLA protein.";		
RL	Nucleic Acids Res. 18:1909-1909(1990).		
RN	[2]		
RN	SEQUENCE FROM N.A.		
RX	MEDLINE=87004555; PubMed=3019668;		
RA	Celestee A.J., Buecker J.L., Ritz R., Wang E.A., Mooney J.M. ;		
RT	"Isolation of the human gene for bone gla protein utilizing mouse and rat cDNA clones.";		
RL	EMBO J. 5:1885-1890(1986).		
RN	[3]		
RP	SEQUENCE OF 52-100.		
RX	MEDLINE=81006914; Pubmed=6967872;		
RA	Poser J.W., Esch F.S., Ling N.C., Price P.A. ;		
RT	"Isolation and sequence of the vitamin K-dependent protein from human bone. Undercarboxylation of the first glutamic acid residue.";		
RL	J. Biol. Chem. 255:8685-8691(1980).		
CC	- I - FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS STRONGLY TO APATITE AND CALCIUM.		
CC	- I - PMM: GAMMA-CARBOXYLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE BINDING OF CALCIN.		
CC	- I - SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X53698; CAA37736.1; -		
DR	EMBL; X51699; CAA35996.1; -		
DR	EMBL; X04143; CAA27763.1; -		
DR	PIR; A03301; GEHU.		
DR	PIR; C25471; C25471.		
DR	PIR; SI2652; SI2652.		
DR	Genew; HGNC:1043; BGLAP.		
DR	MIM: 112260; -		
DR	InterPro: IPRO002384; GLA_bone.		
DR	InterPro: IPRO00294; VitK_dep_GLA.		
DR	pIam; PF00594; gla; 1.		
DR	PRINTS; PR00002; GLABONE.		
DR	SMART; SM00069; GLA; 1.		
DR	PROSITE; PS00011; GUU_CARBOXYLATION; 1.		
KW	Calcium-binding; Gamma-carboxylglutamic acid; Vitamin K; Bone; Signal.		
FT	SIGNAL 1 23 PROBABLE.		
FT	PROPEP 24 51		

FT	CHAIN	52	100	OSTEOCALCIN.
FT	MOD_RES	68	68	GAMMA-CARBOXYGLUTAMIC ACID (IN 9% OF THE MOLECULES).
FT	MOD_RES	72	72	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	75	75	GAMMA-CARBOXYGLUTAMIC ACID.
FT	DISULFID	74	80	GAMMA-CARBOXYGLUTAMIC ACID.
FT	CONFLICT	33	34	MISSING (IN REF. 2).
SO	SEQUENCE	100 AA;	10962 MW;	4DP2A0A80849CB71 CRC64;
Query Match				
Best Local Similarity		51.0%;	Score 25;	DB 1; Length 100;
Matches 25;		Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	25	LNPDCDELADHIGFOEAYRFGYPV	49	
DB	76	LNPDCDELADHIGFOEAYRFGYPV	100	
RESULT 7				
OSTC	FELCA	STANDARD;	PRT;	49 AA.
ID	OSTC_FELCA	STANDARD;	PRT;	49 AA.
AC	P02821;			
DT	21-JUL-1986	(Rel. 01, Created)		
DT	21-JUL-1986	(Rel. 01, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).			
DE	BGLAP.			
GN	Felis silvestris catus (Cat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.			
OX	NCBI-TaxID=9685;			
RN	(1)			
RP	SEQUENCE.			
RX	MEDLINE=85054706; PubMed=6334077;			
RA	Shimomura H., Kanai Y., Sanada K.;			
RT	"Primary structure of cat osteocalcin."			
RL	J. Biochem. 96:405-411(1984).			
CC	-1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS STRONGLY TO APATITE AND CALCIUM.			
CC	-1- PTH: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE BINDING OF CALCIUM.			
CC	-1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN FAMILY.			
CC	PIR: A03304; GECT.			
DR	InterPro: IPR002384; GLA_bone.			
DR	InterPro: IPR00294; VITK_dep_GLA.			
DR	Pfam: PF00594; gla; 1.			
DR	PRINTS: PRO0002; GLABONE.			
DR	SMART: SM00069; GLA: 1.			
DR	PROSITE: PS00011; GLU-CARBOXYLATION; 1.			
KW	Calcium-binding; Gamma-carboxyglutamic acid; vitamin k; Hydroxylation; Bone.			
FT	MOD_RES	9	9	HYDROXYLATION.
FT	MOD_RES	17	17	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	21	21	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	24	24	GAMMA-CARBOXYGLUTAMIC ACID.
FT	DISULFID	23	29	
SO	SEQUENCE	49 AA;	5495 MW;	93D2131PA9F656D3 CRC64;
Query Match				
Best Local Similarity		30.6%;	Score 15;	DB 1; Length 49;
Matches 15;		Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	25	LNPDCDELADHIGFO	39	
DB	25	LNPDCDELADHIGFO	39	
RESULT 8				
OSTC	CANFA	STANDARD;	PRT;	49 AA.
ID	OSTC_CANFA	STANDARD;	PRT;	49 AA.

AC	P81455; 1998 (Rel. 37, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).
GN	BGLAP.
OS	Canis familiaris (Dog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX	NCBI_TaxID=9615;
RN	[1]
RP	SEQUENCE.
RX	MEDLINE=93318657; PubMed=8101026;
RA	Colombo G., Fantì P., Yao C., Malluche H.H.:
RT	"Isolation and complete amino acid sequence of osteocalcin from canine bone."
RL	J. Bone Miner. Res. 8:733-743(1993).
RC	-I- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS STRONGLY TO APATITE AND CALCIUM.
CC	-I- PMM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE BINDING OF CALCIUM.
CC	-I- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN FAMILY.
DR	InterPro: IPR002384; GLA_bone.
DR	InterPro: IPR000294; VitK_dep_GLA.
DR	Pfam: PF00594; gla_1.
DR	PRINTS: PR00062; GLABONE.
DR	SMART: SM00069; GLA_1.
KW	PROSITE: PS00011; GLU_CARBOXYLATION; 1.
KW	Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Hydroxylation; Bone.
FT	MOD_RES 9 9 HYDROXYLATION (BY SIMILARITY).
FT	MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT	DISULFID 23 29
SQ	SEQUENCE 49 AA: 5524 MW: 43121015817CEB6 CRC64;
Query Match	28.6%; Score 14; DB 1; Length 49;
Best Local Similarity	100.0%; Pred. No. 3; Ee -09;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	29 CDELADHGFGEAY 42       Db
DB	29 CDELADHGFGEAY 42
RESULT 9	
OSTC_CHICK	
ID	OSTC_CHICK STANDARD: PRT: 97 AA.
DC	P02822; Q90620;
AC	21-JUL-1986 (Rel. 01, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).
GN	BGLAP.
OS	Gallus gallus (Chicken).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC	Gallus.
OX	NCBI_TaxID=9031;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=White Leghorn; TISSUE=Bone;
RX	MEDLINE=95266465; PubMed=7747623;
RA	Neugebauer B.M., Moore M.A., Broess M., Gerstenfeld L.C., Hauschka P.V.;
RT	"Characterization of structural sequences in the chicken osteocalcin gene: expression of osteocalcin by maturing osteoblasts and by hypertrophic chondrocytes in vitro";

```

RL J. Bone Miner. Res. 10:157-163(1995).
RN [2]
RP SEQUENCE OF 49-97.
RX MEDLINE=82007831; Pubmed=6792200;
RA Carr S.A., Hauschka P.V., Blemann K.;
RT "Gas chromatographic mass spectrometric sequence determination of
RT osteocalcin, a gamma-carboxyglutamic acid-containing protein from
RT chicken bone."
RL J. Biol. Chem. 256:9944-9950(1981).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U10578; AAA78809.1; -.
DR PIR; A03305; GECH.
DR InterPro; IPR002384; GLA_bone.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00594; gla; 1.
DR PRINTS; PR00002; GLABONE.
DR SMART; SM00069; GLA; 1.
DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
KM Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Bone; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 48
FT CHAIN 49 97 OSTEOCALCIN.
FT MOD_RES 65 65 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 72 72 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 71 77 BY SIMILARITY.
SQ SEQUENCE 97 AA; 10707 MW; 768E685C3327D62 CRC64;

Query Match 28.6%; Score 14; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 IGFOEAYRRFGPV 49
Db 84 IGFOEAYRRFGPV 97

RESULT 10
OSTC_RAT
ID OSTC_RAT STANDARD; PRT; 99 AA.
AC P04640;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)
DE (Bone Gla-protein) (BGP).
GN BGLAP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87004555; Pubmed=3019668;
RA Celeste A.J., Buckner J.L., Kriz R., Wang E.A., Wozney J.M.;
RT "Isolation of the human gene for bone gla protein utilizing mouse and
RT rat cDNA clones."
RL EMBO J. 5:1885-1890(1986).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85298305; Pubmed=3875856;
RA Pan L.C., Price P.A.;
RT "The propeptide of rat bone gamma-carboxyglutamic acid protein shares
RT homology with other vitamin K-dependent protein precursors."
RL Proc. Natl. Acad. Sci. U.S.A. 82:6109-6113(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89118266; Pubmed=3265336;
RA Yoon K., Rutledge S.J.C., Buenaga R.F., Rodan G.A.;
RT "Characterization of the rat osteocalcin gene: stimulation of
RT promoter activity by 1,25-dihydroxyvitamin D3."
RL Biochemistry 27:8521-8526(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89251082; Pubmed=2785907;
RA Theofan G., Haberstroth L.M., Price P.A.;
RT "Molecular structure of the rat bone gla protein gene and
RT identification of putative regulatory elements."
RL DNA 8:213-221(1989).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=89145200; Pubmed=2784002;
RA Lian J., Stewart C., Puchacz E., Mackowiak S., Shalhoub V.;
RA Collart D., Zambetti G., Stein G.;
RT "Structure of the rat osteocalcin gene and regulation of vitamin D-
RT dependent expression."
RL Proc. Natl. Acad. Sci. U.S.A. 86:1143-1147(1989).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X04141; CAA27761.1; -.
DR EMBL; M11777; AAA40816.1; -.
DR EMBL; M23637; AAA41761.1; -.
DR EMBL; M25490; AAA53280.1; -.
DR EMBL; J04500; AAA11764.1; -.
DR PIR; A25167; GERT.
DR PIR; A23471; A25471.
DR PIR; A31856; A31856.
DR PIR; A31419; A31419.
DR PIR; A32324; A32324.
DR InterPro; IPR002384; GLA_bone.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00594; gla; 1.
DR PRINTS; PR00002; GLABONE.
DR SMART; SM00069; GLA; 1.
DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
KM Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;
KW Bone; Signal.
FT SIGNAL 1 23 PROBABLE.
FT PROPEP 24 49 PROBABLE.
FT CHAIN 50 99 OSTEOCALCIN.
FT MOD_RES 58 58 HYDROXYLATION.
FT MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 70 70 (BY SIMILARITY).
FT MOD_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 73 73 (BY SIMILARITY).

```

FT DISULFID 72 78 BY SIMILARITY.  
SQ SEQUENCE 99 AA; 10927 MW; 7F18F1866D4E4388 CRC64;  
Query Match 22.4%; Score 11; DB 1; Length 99;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 29 CDELADHIGFQ 39  
| | | | | | | | | |  
DB 78 CDELADHIGFQ 88  
RESULT 11  
ID OSTD\_RABIT STANDARD; PRT; 49 AA.  
AC P33056;  
DT 01-FEB-1995 (Rel. 31, Last Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Osteocalcin (gamma-carboxyglutamic acid-containing protein)  
DE (Bone gla-protein) (BGP).  
GN BGLAP.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_Taxid=9986;  
RN [1]  
RP MEDLINE=92175242; PubMed=1794506;  
RA Viridi A.S., Willis A.C., Hauschka P.V., Triffitt J.T.;  
RT Primary aminoacid sequence of rabbit osteocalcin.;  
RL Biochem. Soc. Trans. 19:3735-3738(1991).  
CC -1 FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
CC STRONGLY TO APATITE AND CALCIUM.  
CC -1 PTH: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
CC BINDING OF CALCIUM.  
CC -1 SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
CC FAMILY.  
DR PIR: A61280; A61280.  
DR InterPro: IPR002384; GLA\_bone.  
DR InterPro: IPR000294; VitK\_dep\_GLA.  
DR Pfam: PF00594; gla; 1.  
DR PRINTS: PR00002; GLABONE.  
DR SMART: SM00069; GLA; 1.  
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.  
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;  
KW Bone.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 9 9 HYDROXYLATION.  
FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
FT DISULFID 23 29 BY SIMILARITY  
SQ SEQUENCE 49 AA; 5431 MW; 7B218871F0312253 CRC64;  
Query Match 20.4%; Score 10; DB 1; Length 49;  
Best Local Similarity 100.0%; Pred. No. 0.00011;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 25 LNPDCDELAD 34  
| | | | | | | | | |  
DB 25 LNPDCDELAD 34  
RESULT 12  
ID NRR3\_HUMAN STANDARD; PRT; 447 AA.  
AC Q13133;  
DT 30-MAY-2000 (Rel. 39, Last Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Oxyterols receptor LXR-alpha (Liver X receptor alpha) (Nuclear orphan

DE receptor LXR-alpha).  
GN NRR13 OR LXR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95262897; PubMed=7744246;  
RA Mangelsdorf D.J.;  
RT "LXR, a nuclear receptor that defines a distinct retinoid response  
RT pathway.";  
RL Genes Dev. 9:1033-1045(1995).  
CC -1 FUNCTION: ORPHAN RECEPTOR. INTERACTION WITH RXR SHIFTS RXR FROM  
CC ITS ROLE AS A STENT DNA-BINDING PARTNER TO AN ACTIVE LIGAND-  
CC BINDING SUBUNIT IN MEDIATING RETINOID RESPONSES THROUGH TARGET  
CC GENES DEFINED BY LXXS. LXXS ARE DNA-TYPE RESPONSE ELEMENTS  
CC CHARACTERIZED BY DIRECT REPEATS OF TWO SIMILAR HEXANUCLEOTIDE HALF-  
CC SITES SPACED BY FOUR NUCLEOTIDES. LXXA PLAYS AN IMPORTANT ROLE IN  
CC THE REGULATION OF CHOLESTEROL HOMEOSTASIS.  
CC -1 SUBUNIT: HETERODIMER OF LXXA AND RXR.  
CC -1 SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1 TISSUE SPECIFICITY: VISCERAL ORGANS SPECIFIC EXPRESSION. STRONG  
CC EXPRESSION WAS FOUND IN LIVER, KIDNEY AND INTESTINE FOLLOWED BY  
CC SPLEEN AND TO A LESSER EXTENT THE ADRENALS.  
CC -1 INDUCTION: BY 9-CIS RETINOIC ACID (9CRA).  
CC -1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
CC NRI SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
CC -----  
DR EMBL: U22662; AAA5856.1; -.  
DR HSSP: P03372; LHCO.  
DR TRANSFAC: T02752; -.  
DR GeneW: HGNC:7966; NRR13.  
DR MIM: 602423; -.  
DR InterPro: IPR000536; Hormone\_rec\_119.  
DR InterPro: IPR001723; Sthrmn\_receptor.  
DR InterPro: IPR001628; znf\_C4steroid.  
DR Pfam: PF00104; hormone\_rec; 1.  
DR Pfam: PF00105; zf-C4; 1.  
DR PRINTS: PR00398; STRDHOMNER.  
DR PRINTS: PR00047; STROIDFINGER.  
DR PRODOM: PD000035; znf\_C4steroid; 1.  
DR SMART: SM00430; HOL1; 1.  
DR SMART: SM00399; znf\_C4; 1.  
DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger.  
FT DNA\_BIND 98 163 NUCLEAR RECEPTOR-TYPE.  
FT ZN\_FING 98 118 C4-TYPE.  
FT ZN\_FING 134 158 C4-TYPE.  
FT DOMAIN 215 434 LIGAND-BINDING (POTENTIAL).  
SQ SEQUENCE 447 AA; 50481 MW; A227B233B0F8C96 CRC64;  
Query Match 14.3%; Score 7; DB 1; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 5 WLGAPVP 11  
| | | | | | | | | |  
DB 4 WLGAPVP 10  
RESULT 13



```

UL84_HCMVA          STANDARD;          PRT;          586 AA.
ID   UL84_HCMVA
AC   P16727;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   65 kDa early nonstructural protein (UL84 protein).
GN   UL84.
OS   Human cytomegalovirus (strain AD169).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
RX   NCBI_TaxID=10360;
RP   MEDLINE=90269039; PubMed=2161319;
RA   Chee M.S., Bankier A.T., Beck S., Bohm R., Brown C.M., Cerny R.,
RA   Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA   Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT   "Analysis of the protein-coding content of the sequence of human
RT   cytomegalovirus strain AD169."
RL   Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: X17403; CAA35358.1; -.
DR   PIR: S09848; WMBEDE.
KW   Nonstructural protein.
FT   DOMAIN          9         19         ARG-RICH (BASIC).
FT   DOMAIN          162        170        LYS-RICH (BASIC).
FT   DOMAIN          171        182        ASP/GLU-RICH (ACIDIC).
SQ   SEQUENCE        586 AA; 65428 MW; 54AB912D6077223F CRC64;

Query Match          14.3%; Score 7; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   42 YRRFYGP 48
      |||||
DB   443 YRRFYGP 449

RESULT 14
UL84_HCMVT
ID   UL84_HCMVT          STANDARD;          PRT;          587 AA.
AC   P29839;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   65 kDa early nonstructural protein (UL84 protein).
GN   UL84.
OS   Human cytomegalovirus (strain Towne).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
RX   NCBI_TaxID=10363;
RP   MEDLINE=92114132; PubMed=1309892;
RA   He Y.S., Xu L., Huang E.S.;
RT   "Characterization of human cytomegalovirus UL84 early gene and
RT   identification of its putative protein product."
RL   J. Virol. 66:1098-1108(1992).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial

```

```

CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: M81432; AAA45947.1; -.
DR   PIR: A41808; WMBETE.
KW   Nonstructural protein.
FT   DOMAIN          9         19         ARG-RICH (BASIC).
FT   DOMAIN          162        170        LYS-RICH (BASIC).
FT   DOMAIN          171        183        ASP/GLU-RICH (ACIDIC).
SQ   SEQUENCE        587 AA; 65388 MW; 13C170E41FB3220B CRC64;

Query Match          14.3%; Score 7; DB 1; Length 587;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   42 YRRFYGP 48
      |||||
DB   444 YRRFYGP 450

RESULT 15
JAK1_BRARE
ID   JAK1_BRARE          STANDARD;          PRT;          1153 AA.
AC   012990; 073880;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Tyrosine-protein kinase Jak1 (EC 2.7.1.112) (Janus kinase 1) (JAK-1).
GN   JAK1.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   RP
RP   FUNCTION.
RP   SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, MUTAGENESIS OF LYS-905, AND
RP   TISSUE=Gastrula;
RX   MEDLINE=97250493; PubMed=9096349;
RA   Conway G., Margolish A., Wong-Madden S., Roberts R.J., Gilbert W.;
RT   "JAK1 kinase is required for cell migrations and anterior
RT   specification in zebrafish embryos."
RL   Proc. Natl. Acad. Sci. U.S.A. 94:3082-3087(1997).
RN   RP
RP   [2]
RP   SEQUENCE OF 443-1153 FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=99445372; PubMed=10515866;
RA   Oates A.C., Brownlie A., Pratt S.J., Irvine D.V., Liao E.C., Paw B.H.,
RA   Dorian K.J., Johnson S.L., Postlewait J.H., Zou L.I., Wilks A.F.;
RT   "Gene duplication of zebrafish JAK2 homologs is accompanied by
RT   divergent embryonic expression patterns: only jak2a is expressed
RT   during erythropoiesis."
RL   Blood 94:2622-2636(1999).
CC   -1- FUNCTION: Tyrosine kinase of the non-receptor type. Appears to be
CC   regulated in early development for specific cell migrations
CC   (epiboly), expression of homeobox protein gooseoid and formation
CC   of anterior structures.
CC   -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC   tyrosine phosphate.
CC   -1- SUBCELLULAR LOCATION: Wholly intracellular, possibly membrane
CC   associated (by similarity).
CC   -1- DEVELOPMENTAL STAGE: Present in the unfertilized egg through to
CC   the blastula stage where it is distributed uniformly. Levels drop
CC   rapidly at four hours development, remain very low until 10 hours,
CC   then gradually increase from 12 hours with a rapid increase at 48
CC   hours. At 48 hours it is concentrated in the region of the gill
CC   arches. Also present in the adult.
CC   -1- DOMAIN: Possesses two phosphotransferase domains. The second one
CC   probably contains the catalytic domain, while the presence of
CC   slight differences suggest a different role for domain 1.
CC   -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. JAK
CC   SUBFAMILY.
CC   -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U82980; AAB54114.1; -
DR EMBL; AF005689; CA006673.1; -
DR HSSP; P11362; 1FGK
DR ZFIN; ZDB-GENE-980526-142; jak1.
DR InterPro; IPR000299; Band_4.1.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00069; pkinase; 2.
DR ProDom; PD000001; Euk_pkinase; 2.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00295; BA1; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_DOM; 2.
DR TRANSFERASE; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW SH2 domain; Repeat; Developmental protein.
FT DOMAIN 335 356 LYS-RICH (BASIC).
FT DOMAIN 435 540 SH2 (ATYPICAL).
FT DOMAIN 580 846 PROTEIN KINASE 1.
FT DOMAIN 872 1150 PROTEIN KINASE 2.
FT NP_BIND 878 886 ATP (BY SIMILARITY).
FT BINDING 905 905 ATP.
FT ACT_SITE 1000 1000 BY SIMILARITY.
FT MOD_RES 1031 1031 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOTIFEN 905 905 K->E: LOSS OF AUTOPHOSPHORYLATION AND
FT CONFLICT 649 649 DEFECTS IN EARLY DEVELOPMENT.
FT CONFLICT 770 770 T -> I (IN REF. 2).
FT CONFLICT 798 798 T -> S (IN REF. 2).
FT CONFLICT 978 978 S -> T (IN REF. 2).
FT CONFLICT 978 978 H -> L (IN REF. 2).
FT CONFLICT 1111 1111 L -> I (IN REF. 2).
SQ SEQUENCE 1153 AA; 132480 MW; 736D5263D03E7450 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -1- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001484; AAD06085.1; -
DR InterPro; IPR001911; Ribosomal_S21.
DR Pfam; PF01165; Ribosomal_S21; 1.
DR PRINTS; PR00976; RIBOSOMAL_S21.
DR ProDom; PD005521; Ribosomal_S21; 1.
DR TIGRPFAMS; TIGR00303; S21P; 1.
DR PROSITE; PS01181; RIBOSOMAL_S21; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 70 AA; 8614 MW; BIDA8696B1A03B7B CRC64;

Query Match 12.2%; Score 6; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

CC -----
DR EMBL: AE000570; AAD07628.1; -
DR TIGR; HP0562; -
DR InterPro; IPR001911; Ribosomal_S21.
DR Pfam; PF01165; Ribosomal_S21; 1.
DR PRINTS; PR00976; RIBOSOMAL_S21.
DR Prodom; PD005521; Ribosomal_S21; 1.
DR TIGRFAMs; TIGR00030; S21P; 1.
DR PROSITE; PS01181; RIBOSOMAL_S21; 1.
KW Ribosomal protein; Complete proteome.
SO SEQUENCE 70 AA; 8613 MW; 417A8696B1A03B76 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 40 EAYRRF 45
      |||||
Db 14 EAYRRF 19

RESULT 18
Y592.CHLPN
ID Y592.CHLPN STANDARD; PRT; 103 AA.
AC Q9Z7W2; Q9K2D2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein CPN0592/CP0156/CPJ0592.
OS CPN0592 OR CP0156 OR CPJ0592.
GN Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99206606; PubMed=10923388;
RX Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=20150255; PubMed=10684935;
RX Ra White T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA Reed T.O., Hickey E.R., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RN pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RN from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SIMILARITY: BELONGS TO THE UP0161 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001643; AAD18731.1; -

```

DR	EMBL; AE0025177; AAF38037.1; ALT_INIT.
DR	EMBL; AP0025647; BAA98799.1; -.
DR	TIGR; CP0156; -.
DR	InterPro; IPR002696; DUF37.
DR	Pfam; PF01809; DUF37; 1.
DR	Prodom; PD004225; DUF37; 1.
DR	TIGRFAMS; TIGR00278; DUF37; 1.
KW	Hypothetical protein; complete proteome.
SQ	SEQUENCE 103 AA; 11751 MW; CE9CA4852EA15A7C CRC64;
Query Match	
Best Local Similarity	12.2%; Score 6; DB 1; Length 103;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 YLYOWL 6 
DB	20 YLYOWL 25
RESULT 19	
ID	CYT_CORJA STANDARD; PRT; 116 AA.
AC	P81061:
DT	15-JUL-1998 (Rel. 36, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DE	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Cystatin (Egg-white cystatin).
OS	Coturnix coturnix japonica [Japanese quail].
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.
OX	NCBI_TaxID=93934;
RN	[1]
RP	SEQUENCE.
RC	TISSUE=Egg white;
RX	MEDLINE=97420480; PubMed=9276465;
RA	Gerhartz B., Engh R.A., Mentele R., Eckerskorn C., Torquato R., Wiltman J., Kolb H.-J., Machleidt W., Fritz H., Auerswald E.A.; "Quail cystatin: isolation and characterisation of a new member of the cystatin family and its hypothetical interaction with cathepsin B.";
RL	FEBS Lett 412:551-558(1997).
-I-	FUNCTION: THIS PROTEIN BINDS TIGHTLY TO AND INHIBITS PAPAIN AND CATHEPSIN B.
CC	-I- SIMILARITY: BELONGS TO THE CYSTATIN FAMILY.
DR	HSSP; P01038; ICEM.
DR	InterPro: IPR000010; Cystatin.
DR	Pfam; PF00031; Cystatin; 1.
DR	SMART; SMO0043; CY; 1.
DR	PROSITE; PS00287; CYSTATIN; 1.
KW	Thiol protease inhibitor; Phosphorylation. FT ACET SITE 9 REACTIVE SITE. FT SITE 53 SECONDARY AREA OF CONTACT. FT DISULEFD 71 81 FT DISULEFD 95 115 FT MOD_RES 80 80 PHOSPHORYLATION. SQ SEQUENCE 116 AA; 13093 MW; 48248621053A2F70 CRC64;
Query Match	
Best Local Similarity	12.2%; Score 6; DB 1; Length 116;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	6 LGAPVP 11 
DB	8 LGAPVP 13
RESULT 20	
ID	MP2_NERDI STANDARD; PRT; 119 AA.
AC	P80255:
DT	01-OCT-1993 (Rel. 27, Created)
DT	01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Non-metallothionein cadmium-binding protein (CD-BP) (MP II).  
 OS Nereis diversicolor (Sandworm) (Hediste diversicolor).  
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;  
 OC Phyllococci; Nereididae; Nereidae.  
 OX NCBI\_TaxID=6352;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=94039032; PubMed=8223553;  
 RT "Amino acid sequence of the small cadmium-binding protein (MP II)  
 RT from Nereis diversicolor (annelida, polychaeta). Evidence for a  
 RT myohemerythrin structure.";  
 RL Eur. J. Biochem. 217:151-156(1993).  
 RN [2]  
 RP SEQUENCE OF 1-33.  
 RX MEDLINE=91347123; PubMed=1908740;  
 RT Demyunck S., Sautiere P., van Beeumen J., Dhainaut-Courtois N.;  
 RT "Homologies between hemerythrins of sipunculids and cadmium-binding  
 RT metalloprotein (MP II) from a polychaete annelid, Nereis  
 RT diversicolor.";  
 RL C. R. Acad. Sci., III, Sci. Vie 312:317-322(1991).  
 CC -1- FUNCTION: MAY ACT AS A BUFFER TO CONTROL THE CONCENTRATION AND  
 CC THEREFORE THE TOXICITY OF CADMIUM.  
 CC -1- SIMILARITY: BELONGS TO THE HEMERYTHRIN FAMILY.  
 CC PIR: S38261; S38261.  
 DR HSP: P02247; 2MR.  
 DR InterPro: IPR002063; Hemerythrin.  
 DR Pfam: PF01814; Hemerythrin; 1.  
 DR PRINTS: PR00186; HEMERYTHRIN.  
 DR PRODOM: PD006099; Hemerythrin; 1.  
 DR TIGRFAMs: TIGR00058; Hemerythrin; 1.  
 DR PROSITE: PS00550; HEMERYTHRINS; 1.  
 KW Metal-binding; Cadmium; Cadmium resistance.  
 SQ SEQUENCE 119 AA; 13454 MW; F78F8DC67F220B02 CRC64;  
 QY 6 LGAPVP 11  
 Db 88 LGAPVP 93  
 Query Match 12.2%; Score 6; DB 1; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 7.7;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RX MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,  
 RA Machleidt W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 RT of cystatin from chicken egg white.";  
 RL Hoppe-Seidler's Z. Physiol. Chem. 364:1487-1496(1983).  
 RN [4]  
 RP CHARACTERIZATION OF PROTEIN.  
 RX MEDLINE=83256421; PubMed=6409085;  
 RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,  
 RA Sumter D.C., Barrett A.J.;  
 RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved  
 RT purification from egg white, characterization, and detection in  
 RT chicken serum.";  
 RL Biochem. J. 211:129-138(1983).  
 RN [5]  
 RP DISULFIDE BONDS.  
 RA Grubb A., Loefberg H., Barrett A.J.;  
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 RT cystatin.";  
 RL FEBS Lett. 170:370-374(1984).  
 RN [6]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=89252033; PubMed=2721673;  
 RA Lader B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,  
 RA Bode W.;  
 RT "The cysteine proteinase inhibitor chicken cystatin is a  
 RT phosphoprotein.";  
 RL FEBS Lett. 248:162-168(1989).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=89052676; PubMed=3191914;  
 RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A.,  
 RA Brzin J., Kos J., Turk V.;  
 RT "The 2.0 A x-ray crystal structure of chicken egg white cystatin and  
 RT its possible mode of interaction with cysteine proteinases.";  
 RL EMBO J. 7:2593-2599(1988).  
 RN [8]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=94087719; PubMed=8263912;  
 RA Diekmann T., Mitschang L., Hofmann M., Kos J., Turk V.,  
 RA Auerwald E.A., Jeanick R., Oschkinat H.;  
 RT "The structures of native phosphorylated chicken cystatin and of a  
 RT recombinant unphosphorylated variant in solution.";  
 RL J. Mol. Biol. 234:1048-1059(1993).  
 CC -1- FUNCTION: THIS PROTEIN BINDS TIGHTLY TO AND INHIBITS A VARIETY OF  
 CC THIOLE PROTEASES INCLUDING FITIN, PAPAEN, AND CATHEPSINS B, C, H,  
 CC AND L. ALTHOUGH ISOLATED FROM EGG WHITE, IT IS ALSO PRESENT IN  
 CC SERUM.  
 CC -1- SIMILARITY: BELONGS TO THE CYSTATIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: J05077; AAA48744.1; -  
 DR PIR: A01274; UDCB;  
 DR PIR: A34456; A34456.  
 DR PDB: 1CEW; 31-JAN-94.  
 DR PDB: 1A67; 27-MAY-98.  
 DR PDB: 1A90; 17-JUN-98.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART: SM00043; Cy; 1.  
 DR PROSITE: PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 139 CYSTATIN.  
 FT ACT\_SITE 32 32 REACTIVE SITE.

```

FT SITE 76 80 SECONDARY AREA OF CONTACT.
FT DISULFID 94 104
FT DISULFID 118 138
FT MOD_RES 103 103 PHOSPHORYLATION (PARTIAL).
FT TURND 35 36
FT TURND 39 40
FT HELIX 42 51
FT TURND 52 52
FT HELIX 53 56
FT TURND 57 58
FT STRAND 63 77
FT TURND 81 95
FT TURND 96 97
FT TURND 99 100
FT HELIX 101 108
FT STRAND 115 125
FT TURND 126 129
FT STRAND 130 138
SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37691 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
DB 31 LGAPVP 36

RESULT 22
YKHE_CAEEL STANDARD; PRT; 171 AA.
AC P34274;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C02C2.6 in chromosome III.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson R., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Garner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
RA Wohlman P.;
RT 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.;
RL Nature 368:32-38(1994).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: L23649; AAA27907.1; -.
CC DR PIR: S44737; S44737.
CC DR Wormpep: C02C2.6; CE00032.
CC DR Hypothetical protein.

```

```

SQ SEQUENCE 171 AA; 20083 MW; 6EE3DF3DEBBD9B5 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 DELADH 35
DB 109 DELADH 114

RESULT 23
NADA_METTH STANDARD; PRT; 304 AA.
AC 027855;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Quinolinate synthetase A.
GN NADA OR MTH1827.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadefora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- FUNCTION: Catalyzes the condensation of iminoaspartate with
CC dihydroxyacetone phosphate to form quinolinate (By similarity).
CC -!- PATHWAY: NAD biosynthesis; aspartate to NAMN; second step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE QUINOLINATE SYNTHETASE A FAMILY.
CC SUBFAMILY 2.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: AE000936; AAB86293.1; -.
CC DR Interpro: IPR003473; NADA.
CC DR Pfam: PF02445; NADA; 1.
CC DR TIGRfams: TIGR00550; nada; 1.
CC DR Pyridine nucleotide biosynthesis: Complete proteome.
CC PYRIDINE NUCLEOTIDE BIOSYNTHESIS: Complete proteome.
CC SEQUENCE 304 AA; 34393 MW; B5C48ACE482143DD CRC64;

Query Match 12.2%; Score 6; DB 1; Length 304;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36
DB 208 ELADHI 213

RESULT 24
YOJM_BACSU STANDARD; PRT; 338 AA.
AC P54550;

```

01-OCT-1996 (Rel. 34, Created)  
 01-OCT-1996 (Rel. 34, Last sequence update)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable NADH-dependent flavin oxidoreductase yqjM (EC 1.-.-.-).  
 GN yqjM  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / JH642;  
 RX MEDLINE=97124195; PubMed=8696508;  
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M., Kobayashi Y.;  
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the Bacillus subtilis genome containing the skin element and many sporulation genes."  
 RL Microbiology 142:3103-3111(1996).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borchert S., Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A., Denzot F., Devigne K.M., Dusterhoff A., Ehrlich S.D., Emerson P.T., Ertlan K.D., Errington J., Fabret C., Ferrati E., Fougere D., Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A., Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klier-Bianchard M., Klein C., Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapdus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetlelle D., Portolillo S., Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S., Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weltenegeger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
 RA The complete genome sequence of the Gram-positive bacterium Bacillus subtilis."  
 RL Nature 390:249-256(1997).  
 RT -1 SIMILARITY: TO THE OYE FAMILY OF NADH-DEPENDENT FLAVIN OXIDOREDUCTASES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: D84432; BAA12619.1; -  
 DR EMBL: Z99116; CAB14314.1; -  
 DR Subtilist; BG11742; yqjM  
 DR InterPro: IPR001153; Oxidored\_FMN.  
 DR Pfam: PF00724; Oxidored\_FMN.  
 KW Hypothetical protein; Oxidoreductase; NAD; FAD; Flavoprotein;  
 KM Complete proteome.  
 SQ SEQUENCE 338 AA; 37583 MW; BDC52D34236326FE CRC64;

Query Match 12.2%; Score 6; DB 1; Length 338;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 33 ADHIF 38  
 DB 229 ADHIF 234  
 RESULT 25  
 Y797\_METJA STANDARD; PRT; 367 AA.  
 ID Y797\_METJA  
 AC Q58207;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein M0797.  
 GN M0797.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D., Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I., Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A., Scott J.L., Geoghegan N.S.M., Weidman J.F., Fumman J.L., Nguyen D., Urtrebeck T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C., Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M., Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii."  
 RL Science 273:1058-1073(1996).  
 RT -1 SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1 SIMILARITY: TO M.JANNASCHII M01507.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U67524; AAB98792.1; -  
 DR TIGR: M0797; -  
 DR InterPro: IPR003838; DUF214.  
 DR Pfam: PF02687; DUF214; 1.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 18 38 POTENTIAL.  
 FT TRANSMEM 239 259 POTENTIAL.  
 FT TRANSMEM 236 316 POTENTIAL.  
 FT TRANSMEM 329 349 POTENTIAL.  
 SQ SEQUENCE 367 AA; 41038 MW; 3CC882FE21ED3746 CRC64;  
 Query Match 12.2%; Score 6; DB 1; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 42 YRRFYG 47  
 DB 178 YRRFYG 183  
 RESULT 26  
 B4G1\_MOUSE STANDARD; PRT; 399 AA.  
 ID B4G1\_MOUSE  
 AC P15535;  
 DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Beta-1,4-galactosyltransferase 1 (EC 2.4.1.1-) (Beta-1,4-galactase 1)  
DE (betaGal-T1) (D4G41-T1) (UDP-galactose:beta-N-acetylglucosamine beta  
DE 1,4-galactosyltransferase 1) (UDP-gal:beta-GlcNAc beta-1,4-  
DE galactosyltransferase 1) [Includes: Lactose synthase A protein  
DE EC 2.4.1.22]; N-acetylglucosamine synthase (EC 2.4.1.90) (Nal  
DE synthetase); Beta-N-acetylglucosaminyl-glycopeptide beta-1,4-  
DE galactosyltransferase (EC 2.4.1.38); Beta-N-acetylglucosaminyl-  
DE glycolipid beta-1,4-galactosyltransferase (EC 2.4.1.1-)].  
GN B4G4L1 OR GGT82 OR GGTB.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Choriata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RX MEDLINE=88273147; PubMed=31434348;  
RA Shaper N.L., Hollis G.F., Douglas J.G., Kirsch I.R., Shaper J.H.;  
RT "Characterization of the full length cDNA for murine beta-1,4-  
RT galactosyltransferase. Novel features at the 5'-end predict two  
RT transational start sites at two in-frame AUGs.";  
RL J. Biol. Chem. 263:10420-10428(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90138913; PubMed=1689054;  
RA Shaper N.L., Wright W.W., Sharper J.H.;  
RT "Murine beta 1,4-galactosyltransferase: both the amounts and  
RT structure of the mRNA are regulated during spermatogenesis.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:791-795(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89350913; PubMed=2504153;  
RA Hollis G.F., Douglas J.G., Shaper N.L., Shaper J.H.,  
RA Stafford-Hollis J.M., Evans R.J., Kirsch I.R.,  
RT "Genomic structure of murine beta-1,4-galactosyltransferase.";  
RL Biochem. Biophys. Res. Commun. 162:1069-1075(1989).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89033997; PubMed=3141392;  
RA Nakazawa K., Ando T., Kimura T., Nishimatsu H.;  
RT "Cloning and sequencing of a full-length cDNA of mouse N-  
RT acetylglucosamine (beta 1-4)galactosyltransferase.";  
RL J. Biochem. 104:165-168(1988).  
RN [5]  
RP SEQUENCE OF 1-63 FROM N.A.  
RX MEDLINE=89207607; PubMed=3149531;  
RA Shaper J.H., Hollis G.F., Shaper N.L.;  
RT "Evidence for two forms of murine beta-1,4-galactosyltransferase  
RT based on cloning studies.";  
RL Biochimie 70:1683-1688(1988).  
RN [6]  
RP SEQUENCE OF 1-20 FROM N.A.  
RX STRAIN-BALB/C;  
RC MEDLINE=93043838; PubMed=1384819;  
RA Harduin-lepers A., Shaper N.L., Mahoney J.A., Shaper J.H.;  
RT "Murine beta 1,4-galactosyltransferase: round spermatid transcripts  
RT are characterized by an extended 5'-untranslated region.";  
RL Glycobiology 2:361-368(1992).  
RN [7]  
RP SEQUENCE OF 1-20 FROM N.A.  
RX STRAIN-BALB/C;  
RC MEDLINE=93300832; PubMed=8314795;  
RA Harduin-lepers A., Shaper J.H., Shaper N.L.;  
RT "Characterization of two cis-regulatory regions in the murine beta  
RT 1,4-galactosyltransferase gene. Evidence for a negative regulatory  
RT element that controls initiation at the proximal site.";  
RL J. Biol. Chem. 268:14348-14359(1993).  
CC -1- FUNCTION: THE GOLGI COMPLEX FORM CATALYZES THE PRODUCTION OF  
CC LACTOSE IN THE LACTATING MAMMARY GLAND AND COULD ALSO BE  
CC RESPONSIBLE FOR THE SYNTHESIS OF COMPLEX-TYPE N-LINKED  
CC OLIGOSACCHARIDES IN MANY GLYCOPROTEINS AS WELL AS THE CARBOHYDRATE

CC		MOTIFIES OF GLYCOLIPIDS.
CC	-I-	FUNCTION: THE CELL SURFACE FORM FUNCTIONS AS A RECOGNITION MOLECULE DURING A VARIETY OF CELL TO CELL AND CELL TO MATRIX INTERACTIONS, AS THOSE OCCURRING DURING DEVELOPMENT AND EGG FERTILIZATION, BY BINDING TO SPECIFIC OLIGOSACCHARIDE LIGANDS ON OPPOSING CELLS OR IN THE EXTRACELLULAR MATRIX.
CC	-I-	CATALYTIC ACTIVITY: UDP-galactose + D-glucose = UDP + lactose.
CC	-I-	CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-glucosaminylglycopeptide = UDP + beta-D-glucosaminylglycopeptide.
CC	-I-	CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-glucosaminylglycopeptide + N-acetyllactosamine = UDP + N-acetyllactosamine.
CC	-I-	PATHWAY: Glycosylation.
CC	-I-	SUBUNIT: HOMODIMER; AND HETERODIMER WITH ALPHA-LACTABULMIN TO FORM LACTOSE SYNTHASE.
CC	-I-	SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORMS IN TRANS CISTERNAE OF GOLGI (LONG AND SHORT FORM) AND ON CELL SURFACE (LONG FORM); SOLUBLE FORM IN BODY FLUIDS.
CC	-I-	ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG/CELL SURFACE FORM (SHOWN HERE) AND A SHORT/GOLGI COMPLEX FORM; ARE PRODUCED BY ALTERNATIVE INITIATION.
CC	-I-	PFM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORMS BY PROTEOLYTIC PROCESSING.
CC	-I-	SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 7.
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
DR	EMBL:	J03880: AAA37297.1; -
DR	EMBL:	D00314: BAA00216.1; -
DR	EMBL:	M27922: AAA58745.1; JOINED.
DR	EMBL:	M27917: AAA58745.1; JOINED.
DR	EMBL:	M27918: AAA58745.1; JOINED.
DR	EMBL:	M27919: AAA58745.1; JOINED.
DR	EMBL:	M27920: AAA58745.1; JOINED.
DR	EMBL:	M27921: AAA58745.1; JOINED.
DR	EMBL:	M27922: AAA58744.1; -
DR	EMBL:	M27917: AAA58744.1; JOINED.
DR	EMBL:	M27918: AAA58744.1; JOINED.
DR	EMBL:	M27919: AAA58744.1; JOINED.
DR	EMBL:	M27920: AAA58744.1; JOINED.
DR	EMBL:	M27921: AAA58744.1; JOINED.
DR	EMBL:	M36289: AAA37294.1; -
DR	EMBL:	L16840: AAA62340.1; -
DR	PIR:	A33396: A33396.
DR	MGD:	MGI:95705: A3396t1.
DR	InterPro:	IPR003859: Galactosyl_T_2.
DR	Pfam:	PF02709: Galactosyl_T_2.1.
KW	Transferase:	Glycosyltransferase; Glycoprotein; Transmembrane;
KW	Signal-anchor:	Golgi stack; Multigene family; Alternative initiation.
FT	CHAIN	1 399
FT	CHAIN	14 399
FT	CHAIN	? 399
FT	INTL_MET	14 14
FT	DOMAIN	1 24
FT	TRANSMEM	25 44
FT	DOMAIN	45 399
FT	CARBOHYD	113 113
FT	DISULFID	131 244
SO	SEQUENCE	399 AA: 44411 MW: 084E3437115f4BDD CNC64;
Query Match		12.2%; Score 6; DB 1; Length 399;
Best Local Similarity		100.0%; Pred No. 22;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 IGFQEA 41  
|||||  
Db 231 IGFQEA 236

RESULT 27  
ID PURK\_CORAM STANDARD; PRT; 413 AA.  
AC 044678;  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DE Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)  
DE (AIR carboxylase) (AIRC).  
GN PURK.  
OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).  
OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteriaceae;  
OC Actinomycetales; Corynebacteriales; Corynebacteriaceae;  
OC Corynebacterium.  
OX NCBI\_TaxID=1697;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 6872;  
RX MEDLINE=96236888; PubMed=8998996;  
RA Chung S.O., Lee J.H., Lee S.Y., Lee D.S.;  
RT "Genomic organization of purK and purE in Brevibacterium ammoniagenes  
ATCC 6872: pure locus provides a clue for genomic evolution.",  
RL FEMS Microbiol. Lett. 137:265-268(1996).  
CC -1- FUNCTION: POSSESSES AN ATPASE ACTIVITY THAT IS DEPENDENT ON THE  
CC PRESENCE OF AIR (AMINOIMIDAZOLE RIBONUCLEOTIDE). THE ASSOCIATION  
CC OF PURK AND PURE PRODUCES AN ENZYME COMPLEX CAPABLE OF CONVERTING  
CC AIR TO CAIR EFFICIENTLY UNDER PHYSIOLOGICAL CONDITION  
CC (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: 5-amino-1-(5-phospho-D-riboseyl)imidazole-4-  
CC carboxylate = 5-amino-1-(5-phospho-D-riboseyl)imidazole + CO(2).  
CC -1- PATHWAY: De novo purine biosynthesis; sixth step.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE PURK / PURT FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X91189; CA62598.1; -.  
DR HSSP: P09029; 1B6R.  
DR InterPro: IPR003135; ATP-grasp.  
DR Pfam: PF02222; ATP-grasp.1.  
DR TIGRFAMs: TIGR01161; purK.1.  
KW purine biosynthesis; lysase; Decarboxylase.  
SQ SEQUENCE 413 AA; 44110 MW; F738B230DF5D119A CRC64;

Query Match 12.2%; Score 6; DB 1; Length 413;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
|||||  
Db 132 LGAPVP 137

RESULT 28  
ID CLPX\_ANASP STANDARD; PRT; 445 AA.  
AC 08YOX7;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE ATP-dependent Clp protease ATP-binding subunit clpX.  
GN CLPX OR CLPC OR ALR3684.  
OS Anabaena sp. (strain PCC 7120).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kurita T., Sasamoto S.,  
RA Matanabe A., Iriuchihara M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
cyanobacterium Anabaena sp. strain PCC 7120.",  
RL DNA Res. 8:205-213(2001).  
CC -1- FUNCTION: ATP-dependent specificity component of the Clp protease.  
CC It directs the protease to specific substrates. Can perform  
CC chaperone functions in the absence of clpP (by similarity).  
CC -1- SUBUNIT: Heterodimer of clpP and clpX (by similarity).  
CC -1- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: AP003593; BAB75383.1; -.  
DR InterPro: IPR003593; AAA\_ATPase.  
DR InterPro: IPR003959; AAA\_ATPase-centr.  
DR InterPro: IPR004487; CLPX.  
DR Pfam: PF00004; AAA.1.  
DR SMART: SM00382; AAA.1.  
DR TIGRFAMs: TIGR00382; clpX.1.  
KW Chaperone; ATP-binding; Zinc-finger; Complete proteome.  
FT ZN\_FING 10 35  
FT NP\_BIND 141 148  
SQ SEQUENCE 445 AA; 48939 MW; C1E021D7831FDBA2 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 445;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 QEAYRR 44  
|||||  
Db 381 QEAYRR 366

RESULT 29  
ID HMGS\_ARATH STANDARD; PRT; 461 AA.  
AC P54873; Q9S707;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE (3-hydroxy-3-methylglutaryl-coenzyme A synthase).  
GN HMGS OR MVAL OR AT4G11820 OR T26M18.30.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Landsberg erecta;  
RX MEDLINE=96144274; PubMed=8566777;  
RA Montanari F., Guillon M., Karst F., Delrot S.;  
RT "Isolation and characterization of a cDNA encoding Arabidopsis  
thaliana 3-hydroxy-3-methylglutaryl-coenzyme A synthase.";



Gene 167:197-201(1995).  
 [2]  
 RC SEQUENCE FROM N.A.  
 RN Connolly E.L., Learned R.M.:  
 RT "Post-transcriptional regulation of HMG-CoA synthase expression in  
 RL Arabidopsis thaliana.";  
 RN Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-cv. Columbia;  
 RX MEDLINE-20083488; PubMed-10617198;  
 RA Mayer K.F.X., Schueler C., Wandut R., Murphy G., Volckaert G.,  
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,  
 RA Harris B., Ansong W., Brandt V., Griehl L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermayer B., Macho R., Mueller M.,  
 RA Kreis M., Delenly M., Palgdomenech P., Watson M., Schmidheini T.,  
 RA Reicher B., Fortetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hohnel J., Zimmermann W., Wedler H., Ridley P.,  
 RA Langham S.A., McCullagh B., Bilham L., Robben J.,  
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,  
 RA Braeken M., Melijens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
 RA Weizenecker T., Bothe G., Kampsperger U., Hilbert H., Braun M.,  
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
 RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koeltter P.,  
 RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Buysbaert C., Gielen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,  
 RA Petrett A., Rajandream M.A., Lyne M., Benes V., Reckmann S.,  
 RA Boriva A., Bloeker H., Scharte M., Grimm M., Loehner T.-H.,  
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fartman B., Granderath K., Danner D., Herzl A.,  
 RA Neumann S., Argitrou A., Vitale D., Liquri R., Piravandi E.,  
 RA Masenat O., Ougley F., Clabaud G., Muendlein A., Felber R.,  
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,  
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
 RA Zaccaria P., Devan M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parrell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Tharaidh J.,  
 RA Stoecking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
 RA Mink P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
 RA Kramer J., Fulton L., Kerdas E., Dante M., Pepin K., Hillier L.,  
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,  
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
 RA Antonolu B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,  
 RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Metero A., Shah R.,  
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
 RA Granaat S., Shoddy N., Hasegawa A., Hamed A., Lodi M., Johnson A.,  
 RA Chen E., Maria M., Martensen R., McCombie W.R.:  
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 402:769-777(1999).  
 CC -1- FUNCTION: THIS ENZYME CONDENSES ACETYL-COA WITH ACETOACETYL-COA TO  
 CC FORM HMG-COA, WHICH IS THE SUBSTRATE FOR HMG-COA REDUCTASE (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxy-3-methylglutaryl-CoA + CoA =  
 CC acetyl-CoA + H(2)O + acetoacetyl-CoA.  
 CC -1- PATHWAY: PRODUCTION OF MEVALONATE FROM HMG-COA PRIOR TO THE  
 CC SYNTHESIS OF STEROIDS AND ISOPRENIDS.  
 CC -1- SIMILARITY: BELONGS TO THE HMG-COA SYNTHASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: X83882; CAA58763.1; -  
 DR EMBL: U79160; AAD00297.1; -  
 DR EMBL: U79161; AAD00298.1; -  
 DR EMBL: AL078606; CAB44320.1; -  
 DR EMBL: AL161532; CAB78225.1; -  
 DR InterPro: IPR000590; HMG-CoA\_synth.  
 DR Pfam: PF01154; HMG-CoA\_synth.1.  
 DR PROSITE: PS01226; HMG-CoA\_SYNTHASE.1.  
 KW Lyase; Sterol biosynthesis.  
 FT ACT\_SITE 117 117 POTENTIAL.  
 FT FT 306 A -> S (IN REF. 1).  
 FT CONFLICT 342 342 K -> N (IN REF. 1).  
 SQ SEQUENCE 461 AA; 51094 MW; F44908369AFC5A8 CRC64;  
 Query Match 12.2%; Score 6; DB 1; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 42 YRRFYG 47  
 DB 444 YRRFYG 449  
 RESULT 30  
 ID SYE\_YERPE STANDARD: PRT: 471 AA.  
 AC Q8ZCK0;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)  
 DE (GLURS)  
 GN GLTX OR YPO2984.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_TaxID-632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CO-92 / Blovay Orientalis;  
 RX MEDLINE-21470413; PubMed-11586360;  
 RA Parhill J., Wren B.W., Thomson N.R., Thibault R.W., Holden M.T.G.,  
 RA Prentice M.B., Sepahnia M., James K.D., Churchill C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Farraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Feltham T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skellon J., Stevens K., Whitehead S., Barrett B.G.:  
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";  
 RL Nature 413:523-527(2001).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +  
 CC diphosphate + L-glutamyl-tRNA(Glu).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 FT Complete proteome.  
 FT SITE 9 "HIGH" REGION.  
 FT SITE 237 241 "KMSKS" REGION.  
 FT BINDING 240 240 ATP (BY SIMILARITY).  
 SQ SEQUENCE 471 AA; 53110 MM; 6176293F008B338B C6C64;  
 Query Match 12.2%; Score 6; DB 1; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 LGAPVP 11  
 DB 217 LGAPVP 222

RESULT 31  
 ICPO\_HSVBR STANDARD: PRT; 532 AA.  
 ID ICPO\_HSVBR  
 AC P28990;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Trans-acting transcriptional protein ICPO.  
 GN 63.  
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicelloviruses.  
 OX NCBI\_TaxID=31520;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9229556; PubMed=1318606;  
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;  
 RT "The DNA sequence of equine herpesvirus-1.";  
 RL Virology 189:304-316(1992).  
 [2]  
 RN [2]  
 RP SEQUENCE BY NMR OF 1-63.  
 RX MEDLINE=94087718; PubMed=8263911;  
 RA Everett R.D., Barlow P.N., Milner A., Orr A., Hope G.,  
 RT Lyon D.;  
 RT "A novel arrangement of zinc-binding residues and secondary structure  
 in the CHC4 motif of an alpha herpes virus protein family.";  
 RL J. Mol. Biol. 234:1038-1047(1993).  
 [3]  
 RN [3]  
 RP STRUCTURE BY NMR OF 1-63.  
 RX MEDLINE=94172642; PubMed=8126734;  
 RA Barlow P.N., Luist B., Milner A., Elliott M., Everett R.D.;  
 RT "Structure of the CHC4 domain by 1H-nuclear magnetic resonance  
 spectroscopy. A new structural class of zinc-finger.";  
 RL J. Mol. Biol. 237:201-211(1994).  
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ICPO PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M86664; AAB02498.1; -  
 DR PIR; I36801; WZBEF5.  
 DR PDB; 1CHC; 30-APR-94.  
 DR InterPro; IPR001841; Znf\_fing.  
 DR Pfam; PF00097; Zf-C3HC4; 1.  
 DR SMART; SMO0184; RING; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Transcription regulation; Trans-acting factor; Activator; Zinc-finger;  
 KW DNA-binding; 3d-structure.  
 FT ZN\_FING 8 47 RING-TYPE.  
 FT METAL 8 47 ZINC 1.

FT METAL 11 11 ZINC 1.  
 FT METAL 24 24 ZINC 2.  
 FT METAL 26 26 ZINC 2.  
 FT METAL 29 29 ZINC 1.  
 FT METAL 32 32 ZINC 1.  
 FT METAL 43 43 ZINC 2.  
 FT METAL 46 46 ZINC 2.  
 FT DOMAIN 210 217 POLY-SER.  
 FT STRAND 19 21  
 FT TURN 22 25  
 FT STRAND 26 28  
 FT TURN 30 31  
 FT HELIX 32 39  
 FT TURN 43 46  
 FT STRAND 53 54  
 SQ SEQUENCE 532 AA; 58629 MM; BACB7E16FA26FDFA C6C64;  
 Query Match 12.2%; Score 6; DB 1; Length 532;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 9 PVPYPD 14  
 DB 482 PVPYPD 487

RESULT 32  
 NOEB\_RHIME STANDARD: PRT; 553 AA.  
 ID NOEB\_RHIME  
 AC Q52893;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Nodulation protein noeb.  
 GN NOEB OR RA0417 OR SWA0774.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OS Plasmid pSyma (megaplasmid 1).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 [1]  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=RCR2011 / SU47;  
 RC MEDLINE=96111489; PubMed=8801423;  
 RA Ardourel M., Lortet G., Maillet F., Roche P., Truchet G.,  
 RA Prome J.-C., Rosenberg C.;  
 RT "In Rhizobium meliloti, the operon associated with the nod box n5  
 comprises nodL, noea and noeb, three host-range genes specifically  
 required for the nodulation of particular Medicago species.";  
 RL Mol. Microbiol. 17:687-699(1995).  
 [2]  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=1021;  
 RC MEDLINE=21396509; PubMed=11481432;  
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,  
 RA Batloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,  
 RA Gujral M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,  
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
 RT "Nucleotide sequence and predicted functions of the entire  
 Sinorhizobium meliloti pSyma megaplasmid.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).  
 CC -1- FUNCTION: NOT KNOWN; DOES NOT SEEM TO PARTICIPATE IN NOD FACTOR  
 SYNTHESIS BUT REQUIRED FOR NODULATION ON SOME SPECIFIC MEDICAGO  
 SPECIES SUCH AS M.LITTORALIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: U26430: AAC44092.1: -
DR EMBL: AE007232: AAK65075.1: -
KW TRANSMEM 4 24 TRANSMEMbrane; Complete proteome.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
SQ SEQUENCE 553 AA: 60732 MW: 1296880222955P4 CRC64;

Query Match 12.2% Score 6; DB 1; Length 553;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LADHIG 37
| | | | |
Db 88 LADHIG 93

RESULT 33
TRSL_HCMVA STANDARD; PRT; 788 AA.
ID TRSL_HCMVA
AC P09695;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Hypothetical protein HHLF1.
GN TRSL.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87169717; PubMed=3031311;
RA Weston K., Barrell B.G.;
RT "Sequence of the short unique region, short repeats, and part of the
RT long repeats of human cytomegalovirus.";
RL J. Mol. Biol. 192:177-208(1986).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohm R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RL Curr. Microbiol. Immunol. 154:125-169(1990).
CC -1- SIMILARITY: BELONGS TO THE US22 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: X17403: CAA35269.1: -
DR EMBL: X04650: CAB37121.1: -
DR PTR: C27349: COBEE3.
DR PTR: S09951: S09951.
DR InterPro: IPR003360: US22.
DR Pfam: PF02393: US22; 1.
KW Hypothetical protein.
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 788 AA: 83981 MW: 604869C4472BC7A CRC64;

```

```

Query Match 12.2% Score 6; DB 1; Length 788;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VPYPDP 15
| | | | |
Db 758 VPYPDP 763

RESULT 34
OS94_MOUSE STANDARD; PRT; 838 AA.
ID OS94_MOUSE
AC P48722;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Osmotic stress protein 94 (Heat shock 70-related protein Apg-1).
GN OSP94 OR Apg1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96218151; PubMed=8647834;
RA Kojima R., Randall J., Brenner B.M., Gullans S.R.;
RT "Osmotic stress protein 94 (osp94). A new member of the Hsp110/SSB
RT gene subfamily.";
RL J. Biol. Chem. 271:12327-12332(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-DDY/STD: TISSUE-Testis;
RX MEDLINE=97160564; PubMed=9006898;
RA Kaneko Y., Nishiyama H., Monoguchi K., Higashitsuji H., Kishishita M.,
RA Fujita J.;
RT "A novel hsp110-related gene, apg-1, that is abundantly expressed in
RT the testis responds to a low temperature heat shock rather than the
RT traditional elevated temperatures.";
RL J. Biol. Chem. 272:2640-2645(1997).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS.
CC -1- INDUCTION: BY HYPEROSMOLAR SALT STRESS.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: U23921: AAC52610.1: -
DR EMBL: D49482: BAA08446.1: -
DR MGD: MGI:107422: Osp94.
DR InterPro: IPR001023: Hsp70.
DR Pfam: PF00012: HSP70; 1.
DR ProDom: PD000089; Hsp70; 2.
DR PROSITE: PS00297; HSP70_1; FALSE_NEG.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
KW ATP-binding.
FT CONFLICT 175 176 TA -> HS (IN REF. 2).
FT CONFLICT 221 221 K -> E (IN REF. 2).
FT CONFLICT 279 279 A -> P (IN REF. 2).
FT CONFLICT 308 308 Q -> R (IN REF. 2).
FT CONFLICT 776 776 M -> K (IN REF. 2).
SQ SEQUENCE 838 AA: 94385 MW: B2C021DDA7EAF0B1 CRC64;

Query Match 12.2% Score 6; DB 1; Length 838;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VPYPDP 15

```

```

Db      452 VPRDP 457

|||||
RESULT 35
PPSA_AERPE STANDARD; PRT: 845 AA.
ID PPSA_AERPE
AC OYEC5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphoenolpyruvate synthase (EC 2.7.9.2) (Pyruvate,water dikinase)
DE (PEP synthase).
GN PPSA OR APE0650.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaei A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamizawa M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kusuda N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
CC -I- CATALYTIC ACTIVITY: ATP + pyruvate + H(2)O = AMP +
CC phosphoenolpyruvate + phosphate.
CC -I- PATHWAY: ESSENTIAL STEP IN GLUCOGENESIS WHEN PYRUVATE AND
CC LACTATE ARE USED AS A CARBON SOURCE.
CC -I- SIMILARITY: BELONGS TO THE PEP-UTILIZING ENZYMES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL: AP000060; BAA79621.1;
DR Interpro: IPR000121; PEP_utilizers.
DR Interpro: IPR002192; PPK_N.term.
DR Pfam: PF00391; PEP_utilizers; 1.
DR Pfam: PF01326; PPK_N.term; 1.
DR Pfam: PF02896; PEP_utilizers_C; 1.
DR Prodom: PD000940; PEP_utilizers; 1.
DR PROSITE: PS00370; PEP_ENZYMES_PHOS_SITE; 1.
DR PROSITE: PS00742; PEP_ENZYMES_2; 1.
DR TRANSFERASE; kinase; ATP-binding; Phosphorylation; Complete proteome.
KW MOD_RES 463 463 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 845 AA; 94799 MW; 55C3CD78E2ECB9A5 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 845;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 40 EAYRRF 45
|||||
Db 71 EAYRRF 76

RESULT 36
MSH2_HUMAN STANDARD; PRT: 934 AA.
ID MSH2_HUMAN
AC P43246; O75488;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein Msh2.
GN MSH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94073959; PubMed=8252616;
RA Fishel R., Lescoe M., Rao M., Copeland N.G., Jenkins N.A.,
RA Garber J., Kane M.F., Kolodner R.D.;
RT "The human mutator gene homolog MSH2 and its association with
RT hereditary nonpolyposis colon cancer.";
RL Cell 75:1027-1038(1993).
RN [2]
RN ERRATUM.
RX MEDLINE=94208055; PubMed=8156592;
RA Fishel R., Lescoe M., Rao M., Copeland N.G., Jenkins N.A.,
RA Garber J., Kane M.F., Kolodner R.D.;
RL Cell 77:167-167(1994).
RN [3]
RN SEQUENCE FROM N.A., AND DISEASE.
RX MEDLINE=95229152; PubMed=7713503;
RA Kolodner R.D., Hall N.R., Lipford J., Kane M.F., Rao M.R.S.,
RA Morrison E., Bishop D.T.,
RA Merchant P., Birn P.J., Burn J., Chapman P., Earabino C.,
RT "Structure of the human MSH2 locus and analysis of two Muir-Torres
RT kindreds for msh2 mutations.";
RL Genomics 24:516-526(1994).
RN [4]
RN SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
RN [5]
RN SEQUENCE OF 375-425 FROM N.A.
RC TISSUE=Blood;
RA Corvello C.M., Bevilacqua R.A.U., Rossi B.M., Simpson A.J.G.;
RT "A novel germline mutation at exon 7 of the hMSH2 gene (417 del G) in
RT a large HNPCC Brazilian kindred.";
RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
RN [6]
RN DNA-BINDING.
RX MEDLINE=95007585; PubMed=7923193;
RA Fishel R., Ewel A., Lescoe M.K.;
RT "Purified human MSH2 protein binds to DNA containing mismatched
RT nucleotides.";
RL Cancer Res. 54:5539-5542(1994).
RN [7]
RN IDENTIFICATION OF MSH2 AS MEMBER OF BASC.
RX MEDLINE=20245492; PubMed=10783165;
RA Wang Y., Cortez D., Yazdi P., Neif N., Elledge S.J., Qin J.;
RT "BASC, a super complex of BRCA1-associated proteins involved in the
RT recognition and repair of aberrant DNA structures.";
RL Genes Dev. 14:927-939(2000).
RN [8]
RN REVIEW.
RX MEDLINE=94310688; PubMed=8036718;
RA Jiricny J.;
RT "Colon cancer and DNA repair: have mismatches met their match?";
RL Trends Genet. 10:164-168(1994).
RN [9]
RN REVIEW ON VARIANTS.
RX MEDLINE=97403931; PubMed=9259192;
RA Papadopoulos N., Lindblom A.;
RT "Molecular basis of HNPCC: mutations of MMR genes.";
RL Hum. Mutat. 10:89-99(1997).
RN [10]
RN VARIANTS HNPCC LEO-622 AND TYR-639.
RX MEDLINE=94084796; PubMed=8261515;
RA Leach F.S., Nicolaides N.C., Papadopoulos N., Liu B., Jen J.,
RA Parsons R., Peltomaki P., Sistonen P., Aaltonen L.A.,

```

RA Nystroem-Lahti M., Guan X.-Y., Zhang J., Meltzer P.S., Yu J.-W.,  
 RA Kao F.-T., Chen D.J., Cerosaletti K.M., Fournier R.E.K., Todd S.,  
 RA Lewis T., Leach R.J., Naylor S.L., Welsssbach J., Mecklin J.-P.,  
 RA Jaerinen H., Petersen G.M., Hamilton S.R., Green J., Jass J.,  
 RA Watson P., Lynch H.T., Trent J.M., de la Chapelle A., Kinzler K.W.,  
 RA Vogelstein B.,  
 RT "Mutations of a muts homolog in hereditary nonpolyposis colorectal  
 RT cancer.";  
 RL Cell 75:1215-1225(1993).  
 RN [11]  
 RP VARIANT HNPCC ASN-596 DEL.  
 RX MEDLINE-95179130; PubMed-7874129;  
 RA Mary J.-L., Bishop T., Kolodner R.D., Lipford J.R., Kane M.F.,  
 RA Weber W., Thorhorst J., Mueller H., Spycher R., Scott R.J.;  
 RT "Mutational analysis of the hMSH2 gene reveals a three base pair  
 RT deletion in a family predisposed to colorectal cancer development.";  
 RL Hum. Mol. Genet. 3:2067-2069(1994).  
 RN [12]  
 RP VARIANT HIS-96.  
 RX MEDLINE-95243220; PubMed-7726159;  
 RA Wijnen J., Vasen H., Khan P.M., Menko F.H., van der Klift H.,  
 RA van Leeuwen C., van den Broek M., van Leeuwen-Cornelisse I.,  
 RA Nagengast F., Meijers-Helboer A., Lindhout D., Griffioen G., Cats A.,  
 RA Kleibauer J., Varesco L., Bertario L., Bisgaard M.-L., Mohr J.,  
 RA Fodde R.;  
 RT "Seven new mutations in hMSH2, an HNPCC gene, identified by  
 RT denaturing gradient-gel electrophoresis.";  
 RL Am. J. Hum. Genet. 56:1060-1066(1995).  
 RN [13]  
 RP VARIANT PHE-390 AND LYS-419.  
 RX MEDLINE-9630508; PubMed-8690195;  
 RA Konishi M., Kikuchi-Yanoshita R., Tanaka K., Muraoka M., Onda A.,  
 RA Okumura Y., Kishi N., Iwama T., Mori T., Koike M., Ushio K., Chiba M.,  
 RA Nomizu S., Konishi F., Tsumomiyama J., Miyaki M.;  
 RT "Molecular nature of colon tumors in hereditary nonpolyposis colon  
 RT cancer, familial polyposis, and sporadic colon cancer.";  
 RL Gastroenterology 111:307-317(1996).  
 RN [14]  
 RP VARIANT ASP-322.  
 RX MEDLINE-96163505; PubMed-8566964;  
 RA Mallika Y.K., Chudina A.P., Belev N.F., Alday P., Bochkov N.P.,  
 RA Boersde J.-M.;  
 RT "CPG dinucleotides in the hMSH2 and hMLH1 genes are hotspots for  
 RT HNPCC mutations.";  
 RL Hum. Genet. 97:251-255(1996).  
 RN [15]  
 RP VARIANT HNPCC ASN-596 DEL, AND VARIANT HIS-167.  
 RX MEDLINE-97026284; PubMed-8872463;  
 RA Moslein G., Tester D.J., Lindor N.M., Honchel R., Cunningham J.M.,  
 RA French A.J., Halling K.C., Schwab M., Goretzki P., Thibodeau S.N.;  
 RT "Microsatellite instability and mutation analysis of hMSH2 and hMLH1  
 RT in patients with sporadic, familial and hereditary colorectal  
 RT cancer.";  
 RL Hum. Mol. Genet. 5:1245-1252(1996).  
 RN [16]  
 RP VARIANT CRC TYR-506.  
 RX MEDLINE-96390800; PubMed-8797773;  
 RA Han H.-J., Yuan Y., Ku J.-L., Oh J.-H., Won Y.-J., Kang K.-J.,  
 RA Kim K.-Y., Kim S., Kim C.-Y., Kim J.-P., Oh N.-G., Lee K.-H., Choe K.-J.,  
 RA Nakamura Y., Park J.-G.;  
 RT "Germline mutations of hMLH1 and hMSH2 genes in Korean hereditary  
 RT nonpolyposis colorectal cancer.";  
 RL J. Natl. Cancer Inst. 88:1317-1319(1996).  
 RN [17]  
 RP VARIANT GLN-46.  
 RX MEDLINE-96293410; PubMed-8700523;  
 RA Bubb V.J., Curtis L.J., Cunningham C., Dunlop M.G., Carothers A.D.,  
 RA Morris R.G., White S., Bird C.G., Wyllie A.H.;  
 RT "Microsatellite instability and the role of hMSH2 in sporadic  
 RT colorectal cancer.";  
 RL Oncogene 12:2641-2649(1996).  
 RN [18]  
 RP VARIANTS HNPCC THR-305; THR-834 AND ASN-596 DEL.

RX MEDLINE-97456423; PubMed-9311737;  
 RA Wijnen J., Khan P.M., Vasen H., van der Klift H., Mulder A.,  
 RA van Leeuwen-Cornelisse I., Bakker B., Losekoot M., Moeller P.,  
 RA Fodde R.;  
 RT "Hereditary nonpolyposis colorectal cancer families not complying  
 RT with the Amsterdam criteria show extremely low frequency of  
 RT mismatch-repair-gene mutations.";  
 RL Am. J. Hum. Genet. 61:329-335(1997).  
 RN [19]  
 RP VARIANT HNPCC CYS-323.  
 RX MEDLINE-97362414; PubMed-9240418;  
 RA Akiyama Y., Tsubouchi N., Yuasa Y.;  
 RT "Frequent somatic mutations of hMSH3 with reference to microsatellite  
 RT instability in hereditary nonpolyposis colorectal cancers.";  
 RL Biochem. Biophys. Res. Commun. 236:248-252(1997).  
 RN [20]  
 RP VARIANT SER-596.  
 RX MEDLINE-97147120; PubMed-8993976;  
 RA Viel A., Genuardi M., Capozzi E., Leonardi F., Bellacosa A.,  
 RA Paravatou-Petsotas M., Pomponi M.G., Fornasari M., Percesepe A.,  
 RA Roncucci L., Tamassia M.G., Benatti P., Ponz de Leon M., Valenti A.,  
 RA Covino M., Anli M., Foletto M., Bolocchi M., Neri G.;  
 RT "Characterization of MSH2 and MLH1 mutations in Italian families with  
 RT hereditary nonpolyposis colorectal cancer.";  
 RL Genes Chromosomes Cancer 18:8-18(1997).  
 RN [21]  
 RP VARIANT ASP-322.  
 RX MEDLINE-97242567; PubMed-9087566;  
 RA Wu Y., Nystroem-Lahti M., Ostinga J., Looman M.W.G., Peltomaki P.,  
 RA Aaltonen L.A., de la Chapelle A., Hofstra R.M.W., Buys C.H.C.M.;  
 RT "MSH2 and MLH1 mutations in sporadic replication error-positive  
 RT colorectal carcinoma as assessed by two-dimensional DNA  
 RT electrophoresis.";  
 RL Genes Chromosomes Cancer 18:269-278(1997).  
 RN [22]  
 RP VARIANT HNPCC VAL-562.  
 RX MEDLINE-97201114; PubMed-9048925;  
 RA Beck N.E., Tomlinson I.P.M., Homfray T., Frayling I., Hodgson S.V.,  
 RA Harcopos C.J., Bodmer W.F.;  
 RT "Use of SSCP analysis to identify germline mutations in HNPCC  
 RT families fulfilling the Amsterdam criteria.";  
 RL Hum. Genet. 99:219-224(1997).  
 RN [23]  
 RP VARIANT HNPCC PHE-697, AND VARIANT ASP-322.  
 RX MEDLINE-97442278; PubMed-9298827;  
 RA Wehner M., Buschhausen L., Lambert C., Kruse R., Caspari R.,  
 RA Proping P., Friedl W.;  
 RT "Hereditary nonpolyposis colorectal cancer (HNPCC): eight novel  
 RT germline mutations in hMSH2 or hMLH1 genes.";  
 RL Hum. Mutat. 10:241-244(1997).  
 RL  
 Query Match 12.28; Score 6; DB 1; Length 934;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 7 GAPVPY 12  
 Db 614 GAPVPY 619  
 ID SKIM\_HUMAN STANDARD; PRT; 1246 AA.  
 AC Q15477; Q12902; O15005; Q15476;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Helicase SKI2W (Helicase-like protein) (HLP).  
 GN SKI2VL OR SKI2V OR SKI2W OR DDX13 OR W.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;



RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Pacleb J.M.,  
RA Palazon D., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Relnekt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Slden-Klamos I., Stempson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svrtyskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zheng G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
RA Gdbbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
RN [3]  
RP ALTERNATIVE SPLICING.  
RX MEDLINE=90003228; PubMed=2507168;  
RA Brown N.H., King D.L., Milcox M., Kafatos F.C.;  
RT "Developmentally regulated alternative splicing of *Drosophila* integrin  
RT ps2 alpha transcripts.";  
RL Cell 59:185-195(1989).  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE=98325030; PubMed=9660786;  
RA Graner M.W., Bunch T.A., Baumgartner S., Kerschen A., Brower D.L.;  
RT "Splice variants of the *Drosophila* ps2 integrin differentially  
RT interact with RGD-containing fragments of the extracellular proteins  
RT liggrin, ten-m, and D-laminin 2.";  
RL J. Biol. Chem. 273:18235-18241(1998).  
CC -1- FUNCTION: ALPHA-PS2/BETA-PS IS A RECEPTOR FOR TIGRIN. ALSO BINDS  
CC NO WING BLISTER AND TEN-M  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA  
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A  
CC DISULFIDE BOND. ALPHA-PS2 ASSOCIATES WITH BETA-PS.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: PS2C (SHOWN HERE) AND PS2M8;  
CC MAY BE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- DEVELOPMENTAL STAGE: THE RELATIVE RATIO OF ISOFORM 1/PS2C AND  
CC ISOFORM 2/PS2M8 VARIES WIDELY DURING DEVELOPMENT.  
CC -1- PPM: THE HEAVY-LIGHT CHAIN CLEAVAGE SITE IS EITHER IN 1230-1231,  
CC OR 1233-1234, OR 1243-1244.  
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>).  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AL09059; AAC12788.1; -  
DR EMBL: AE003503; AAF648661.1; -  
DR PIR: A29637; A29637.  
DR HSSP: P11215; 1A8X.  
DR FLYBase; FBgn0001250; 1f.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF01839; FG-GAP; 5.  
DR PRINTS; PRO1185; INTEGRIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Repeat; Alternative splicing.  
FT SIGNAL 1  
FT CHAIN 31  
FT CHAIN 32 1396 POTENTIAL.  
FT CHAIN 32 71243 INTEGRIN ALPHA-PS2.  
FT CHAIN 1244 1396 INTEGRIN ALPHA-PS2 HEAVY CHAIN.  
FT CHAIN 32 1341 INTEGRIN ALPHA-PS2 LIGHT CHAIN.  
FT DOMAIN 1342 1366 EXTRACELLULAR (POTENTIAL).  
FT TRANSSEM 1342 1366 POTENTIAL.  
FT DOMAIN 1367 1396 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 47 115 FG-GAP 1.  
FT REPEAT 128 180 FG-GAP 2.

FT	REPEAT	197	255	FG-GAP 3.
FT	REPEAT	276	327	FG-GAP 4.
FT	REPEAT	329	397	FG-GAP 5.
FT	REPEAT	398	457	FG-GAP 6.
FT	REPEAT	463	517	FG-GAP 7.
FT	DOMAIN	932	938	SER-RICH.
FT	DOMAIN	1015	1023	SER-RICH.
FT	DOMAIN	1049	1054	SER-RICH.
FT	DOMAIN	1217	1226	SER-RICH.
FT	CARBOHYD	69	69	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	209	209	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	322	322	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	584	584	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	598	598	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	741	741	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	783	783	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	833	833	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	959	959	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1005	1005	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1299	1299	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1307	1307	N-LINKED (GLCNAC. . .)
FT	VARSPLIC	225	249	N-LINKED (GLCNAC. . .)
FT	CONFLICT	29	29	MISSING (IN ISOFORM PS2M8).
FT	CONFLICT	968	976	A -> G (IN REF. 1).
FT	CONFLICT	1063	1064	SPROVEQR -> RSQASGATA (IN REF. 1).
FT	CONFLICT	1235	1235	MISSING (IN REF. 1).
FT	CONFLICT	1242	1242	N -> D (IN REF. 1).
FT	CONFLICT	1245	1245	E -> K (IN REF. 1).
FT	CONFLICT	1245	1245	L -> Q (IN REF. 1).
SO	SEQUENCE	1396 AA;	154321 MW;	2384B07DDBA28372 CRC64;
Query Match				
Best Local Similarity		12.28;	Score 6;	DB 1;
Matches		6;	Conservative	0;
			Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	3 YOMIGA 8			
Db	128 YOMIGA 133			
RESULT 39				
PTPD_HUMAN	STANDARD:	PRF:	1912 AA.	
AC	P23468;			
DC	01-NOV-1991 (Rel. 20, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DR	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Protein-tyrosine phosphatase delta precursor (EC 3.1.3.48) (R-PTP-delta).			
GN	PTPRD.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
XM	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-1178.			
RX	MEDLINE=95204468; PubMed=7896816;			
RA	Pulido R., Krueger N.X., Serris-Pages C., Saito H., Streuli M.;			
RT	"Molecular characterization of the human transmembrane protein-tyrosine phosphatase delta. Evidence for tissue-specific expression of alternative human transmembrane protein-tyrosine phosphatase delta isoforms."			
RT	isoforms."			
RL	J. Biol. Chem. 270:6722-6728(1995).			
RN	[2]			
RP	SEQUENCE OF 390-1912 FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=91006018; PubMed=21701019;			
RA	Krueger N.X., Streuli M., Saito H.;			
RT	"Structural diversity and evolution of human receptor-like protein tyrosine phosphatases."			
RL	EMBO J. 9:3241-3252(1990).			
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2O) -> protein tyrosine + phosphate.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			

CC -1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by  
 CC alternative splicing.  
 CC -1- PTM: A CLEAVAGE OCCURS THAT SEPARATES THE EXTRACELLULAR DOMAIN  
 CC FROM THE TRANSMEMBRANE SEGMENT.  
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 8 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L38929; AAC1749.1; -  
 CC EMBL: X54133; CAA38068.1; -  
 CC PIR: S12052; S12052.  
 CC DR HSBP; P18052; 11FO.  
 CC DR HSBP; H6NC; 9668; PTPND.  
 CC DR MIM; 601598; -  
 CC DR InterPro: IPR003961; FN\_III.  
 CC DR InterPro: IPR003962; FNIII\_repeat.  
 CC DR InterPro: IPR003006; Ig\_MHC.  
 CC DR InterPro: IPR003598; Ig\_C2.  
 CC DR InterPro: IPR000387; TYR\_phosphatase.  
 CC DR Pfam: PF00041; fn3; 8.  
 CC DR Pfam: PF00047; Ig; 3.  
 CC DR Pfam: PF00102; Y\_phosphatase; 2.  
 CC DR PRINTS: PRO0014; FNTPFIII.  
 CC DR PRINTS: PRO0070; PTPPHPTASE.  
 CC DR SMART: SM00060; FN3; 8.  
 CC DR SMART: SM00408; IGC2; 3.  
 CC DR SMART: SM00194; PTPC; 2.  
 CC DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 2.  
 CC DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 2.  
 CC DR PROSITE: PS50055; TYR\_PHOSPHATASE\_PTP; 2.  
 CC DR HydroLase; Receptor; Glycoprotein; Signal; Transmembrane; Repeat;  
 CC Immunoglobulin domain; Alternative splicing.  
 CC FT SIGNAL 1 20  
 CC FT CHAIN 1 1912  
 CC FT DOMAIN 21 1265  
 CC FT TRANSMEM 1266 1290  
 CC FT DOMAIN 1291 1912  
 CC FT DOMAIN 23 115  
 CC FT DOMAIN 118 225  
 CC FT DOMAIN 232 318  
 CC FT DOMAIN 320 414  
 CC FT DOMAIN 417 513  
 CC FT DOMAIN 516 606  
 CC FT DOMAIN 609 708  
 CC FT DOMAIN 711 822  
 CC FT DOMAIN 825 916  
 CC FT DOMAIN 918 1017  
 CC FT DOMAIN 1020 1137  
 CC FT DOMAIN 1375 1618  
 CC FT DOMAIN 1619 1912  
 CC FT ACT\_SITE 1553 1553  
 CC FT ACT\_SITE 1844 1844  
 CC FT SITE 1175 1178  
 CC FT CARBOHYD 254 254  
 CC FT CARBOHYD 299 299  
 CC FT CARBOHYD 724 724  
 CC FT CARBOHYD 832 832  
 CC FT CARBOHYD 181 189  
 CC FT VASPLIC 226 229  
 CC FT VASPLIC 775 783  
 CC FT VASPLIC 609 1137  
 CC FT MURAGEN 1178 1178  
 CC SEQUENCE 1912 AA; 214759 MW; 3AEB8BCD32182E26 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 1912;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 31 ELADHI 36  
 Db 1341 ELADHI 1346  
 RESULT 40  
 ID ERY3\_SACER STANDARD; PRT; 3172 AA.  
 AC 003133; 054097; 099270;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Erythronolide synthase, modules 5 and 6 (EC 2.3.1.94) (ORF 3) (6-  
 DE deoxyerythronolide B synthase III) (DEBS 3).  
 GN ERYA  
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;  
 CC Saccharopolyspora.  
 CX NCBI\_Taxid=1836;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL 2338;  
 RA MEDLINE=91043075; PubMed=2234082.  
 RT Cortes J., Haydock S.F., Roberts G.A., Bevilitt D.J., Leadlay P.F.;  
 RT "An unusually large multifunctional polypeptide in the erythromycin-  
 RT producing polyketide synthase of Saccharopolyspora erythraea."  
 RL Nature 348:176-178(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91220065; PubMed=2024119;  
 RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;  
 RT "Modular organization of genes required for complex polyketide  
 RT biosynthesis".  
 RL Science 252:675-679(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL 2338;  
 RX MEDLINE=92155230; PubMed=1740151;  
 RA Bevilitt D.J., Cortes J., Haydock S.F., Leadlay P.F.;  
 RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea.  
 RT Cloning of the structural gene, sequence analysis and inferred domain  
 RT structure of the multifunctional enzyme.";  
 RL Eur. J. Biochem. 204:39-49(1992).  
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-  
 CC deoxyerythronolide B.  
 CC -1- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOTRANSFERINS.  
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN  
 CC BIOSYNTHESIS.  
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH  
 CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3  
 CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH  
 CC SYNTHASE PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS  
 CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,  
 CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,  
 CC RESPECTIVELY.  
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),  
 CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER  
 CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),  
 CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE  
 CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION  
 CC OF THE FULL-LENGTH CHAIN.  
 CC -1- SIMILARITY: TO FATIY ACID SYNTHASE (FAS).  
 CC -1- SIMILARITY: CONTAINS 2 ACTYL CARRIER DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way



CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>  
 CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
 -----  
 CC EMBL; X56107; CAA39583.1; -  
 DR EMBL; X62677; AAA26495.1; -  
 DR EMBL; X62569; CAA44449.1; -  
 DR HSSP; P00101; ICCH  
 DR InterPro; IPR002198; ADH\_short.  
 DR InterPro; IPR001227; Ac\_transferase.  
 DR InterPro; IPR000794; ketoacyl-synt.  
 DR InterPro; IPR003880; Ppanne\_attach.  
 DR InterPro; IPR001031; Thioesterase.  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00109; ketoacyl-synt; 2.  
 DR Pfam; PF00550; pp-binding; 2.  
 DR Pfam; PF00698; Acyl\_transf; 2.  
 DR Pfam; PF00975; Thioesterase; 1.  
 DR Pfam; PF02801; ketoacyl-synt; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.  
 DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 2.  
 DR PROSITE; PS50075; ACP\_DOMAIN; 2.  
 KW Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;  
 KW Phosphopantetheine; Multifunctional enzyme.  
 FT DOMAIN 1 1484  
 FT DOMAIN 1 1484  
 FT DOMAIN 37 484  
 FT DOMAIN 534 878  
 FT DOMAIN 1116 1298  
 FT DOMAIN 1394 1464  
 FT DOMAIN 1488 1954  
 FT DOMAIN 2021 2335  
 FT DOMAIN 2555 2735  
 FT DOMAIN 2821 2891  
 FT DOMAIN 2926 3172  
 FT ACT\_SITE 199 199  
 FT ACT\_SITE 643 643  
 FT NF\_BIND 1118 1164  
 FT BINDING 1427 1427  
 FT ACT\_SITE 1661 1661  
 FT ACT\_SITE 2112 2112  
 FT NF\_BIND 2557 2605  
 FT BINDING 2854 2854  
 FT BINDING 231 231  
 FT BINDING 240 240  
 FT BINDING 289 289  
 FT BINDING 493 493  
 FT BINDING 517 517  
 FT BINDING 510 510  
 FT BINDING 513 513  
 FT BINDING 525 525  
 FT BINDING 536 536  
 FT BINDING 547 551  
 FT BINDING 553 553  
 FT BINDING 673 673  
 FT BINDING 716 716  
 FT BINDING 736 736  
 FT BINDING 896 896  
 FT BINDING 896 896  
 FT BINDING 986 986  
 FT BINDING 986 986  
 FT BINDING 1108 1116  
 FT BINDING 1124 1126  
 FT BINDING 1132 1132  
 FT BINDING 1192 1192  
 FT BINDING 1194 1194  
 FT BINDING 1277 1278  
 FT BINDING 1385 1390  
 FT BINDING 1485 1485  
 FT BINDING 1518 1518  
 FT BINDING 1601 1601  
 FT BINDING 1724 1725  
 FT BINDING 1732 1732  
 O -> L (IN REF. 2).  
 LP -> FA (IN REF. 2).  
 G -> R (IN REF. 2).  
 V -> L (IN REF. 2).  
 A -> R (IN REF. 2).  
 R -> A (IN REF. 2).  
 MISSING (IN REF. 2).  
 AHK -> GIT (IN REF. 2).  
 R -> ROR (IN REF. 2).  
 R -> RELPYRFOROR (IN REF. 1).  
 GVAAPH -> VLSLSD (IN REF. 2).  
 RTMPLLEPLA -> ARTMSWR (IN REF. 2).  
 MISSING (IN REF. 2).  
 L -> V (IN REF. 2).  
 A -> R (IN REF. 2).  
 AA -> RR (IN REF. 2).  
 LCDORE -> STARR (IN REF. 2).  
 MISSING (IN REF. 2).  
 G -> R (IN REF. 2).  
 V -> L (IN REF. 2).  
 LP -> FA (IN REF. 2).  
 O -> L (IN REF. 2).

FT CONFLICT 1739 1743 GPAWG -> ARRA (IN REF. 2).  
 FT CONFLICT 1762 1762 T -> S (IN REF. 2).  
 FT CONFLICT 2252 2252 D -> DGAD (IN REF. 2).  
 FT CONFLICT 2275 2277 OSP -> AVA (IN REF. 2).  
 FT CONFLICT 2408 2408 LA -> GR (IN REF. 2).  
 FT CONFLICT 2420 2421 LA -> S (IN REF. 2).  
 FT CONFLICT 2443 2444 NA -> TH (IN REF. 2).  
 FT CONFLICT 2596 2596 A -> G (IN REF. 2).  
 FT CONFLICT 2609 2609 P -> A (IN REF. 2).  
 FT CONFLICT 2715 2722 RRAEGRRA -> AVRAVRR (IN REF. 1).  
 FT CONFLICT 2754 2754 D -> E (IN REF. 2).  
 SO SEQUENCE 3172 AA; 331474 MW; DBBD5094E7DDDD5F CRC64;  
 Query Match 12.28; Score 6; DB 1; Length 3172;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 32 LADHIG 37  
 DB 2886 LADHIG 2891  
 RESULT 41  
 YBAL\_SCHPO STANDARD; PRT; 69 AA.  
 AC 096VGI;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 GN SPBCL19.18.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_Taxid=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Pettit T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voiclaert G., Aert R., Robben J., Gymnietz B.,  
 RA Welteens J., Vackeert E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesli D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt P., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt G., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shipkovski G.V., Ussery D., Barrell B.G., Nurse P.,  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC -I- SIMILARITY: BELONGS TO THE UPF0203 (15E1.1) FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AL022117; CAC51385.1; -

DR Hypothetical protein.

SO SEQUENCE 69 AA; 7965 MW; 97A3A81350D0105F CRC64;

Query Match 10.2%; Score 5; DB 1; Length 69;

Best Local Similarity 100.0%; Pred. No. 63;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 DCDL 32

DB 37 DCDL 41

RESULT 42

Y066\_HALN1 STANDARD; PRT; 76 AA.

AC 09HS08;

DT 16-OCT-2001 (Rel. 40; Created)

DT 16-OCT-2001 (Rel. 40; Last sequence update)

DT 16-OCT-2001 (Rel. 40; Last annotation update)

DE Hypothetical protein Vng0066h.

GN VNG0066H.

OS Halobacterium sp. (strain NRC-1).

OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;

OC Halobacteriaceae; Halobacterium.

OX NCBI\_TaxID=64091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-20504483; PubMed-11016950;

RA Ng W.V., Kennedy S.P., Maharas G.G., Bergquist B., Pan M., RA Shukla H.D., Laskey S.R., Baliga N.S., Thorsson V., Shrogha J., RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A., RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W., RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H., RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.; RA "Genome sequence of Halobacterium species NRC-1." Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

RL -1- SIMILARITY: BELONGS TO THE UPF0175 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AE004976; AAG18704.1; -

DR InterPro: IPR005368; UPF0175.

DR Pfam: PF03683; UPF0175; 1.

KW Hypothetical protein; Complete proteome.

SO SEQUENCE 76 AA; 8577 MW; 4083B6DFA66C87B CRC64;

Query Match 10.2%; Score 5; DB 1; Length 76;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 DELAD 34

DB 48 DELAD 52

RESULT 43

Y886\_METTH STANDARD; PRT; 92 AA.

AC 026972;

DT 15-JUL-1998 (Rel. 36; Created)

DT 15-JUL-1998 (Rel. 36; Last sequence update)

DT 16-OCT-2001 (Rel. 40; Last annotation update)

DE Hypothetical protein MTH886.

GN MTH886.

OS Methanobacterium thermoautotrophicum.

OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;

OC Methanobacteriaceae; Methanothermobacter.

OX NCBI\_TaxID=187420;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN-Delta H;

RX MEDLINE-98037514; PubMed-9371463;

RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J., RA Alredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K., RA Harrison D., Hoang L., Keagle P., Lunn W., Pothier B., Qiu D., RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R., RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S., RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M., RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.; RA "Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics." J. Bacteriol. 179:7135-7155(1997).

RL -1- SIMILARITY: TO M.JANNASCHIT M0782.1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AE000864; AAB85384.1; -

DR Hypothetical protein; Complete proteome.

SO SEQUENCE 92 AA; 10498 MW; 01214D51FDCFEZB3 CRC64;

Query Match 10.2%; Score 5; DB 1; Length 92;

Best Local Similarity 100.0%; Pred. No. 82;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPV 10

DB 26 LGAPV 30

RESULT 44

NCIM\_HUMAN STANDARD; PRT; 105 AA.

AC 095178;

DT 15-JUL-1999 (Rel. 38; Created)

DT 15-JUL-1999 (Rel. 38; Last sequence update)

DT 15-JUN-2002 (Rel. 41; Last annotation update)

DE NADH-ubiquinone oxidoreductase AGCG subunit, mitochondrial precursor (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-AGCG) (CI-AGCG).

GN NDUFB2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-99097250; PubMed-9878551;

RA Loeffen J.L.C.M., Tiepels R.H., van den Heuvel L., Schuelke M., RA Buskens C.A.F., Smeets R.J.P., Trijbels J.M.F., Smeitink J.A.M.; RA "cDNA of eight nuclear encoded subunits of NADH:ubiquinone RT oxidoreductase: human complex I cDNA characterization completed." Biochem. Biophys. Res. Commun. 253:415-422(1998).

RL [2]

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Blood;

RX MEDLINE-20499367; PubMed-11042152;

RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G., RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W., RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;

```

RT      "Cloning and functional analysis of cDNAs with open reading frames for
RT      300 previously undefined genes expressed in C3H4+ hematopoietic
RT      stem/progenitor cells."
RL      Genome Res. 10:1546-1560(2000).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Lung;
RA      Strausberg R.;
RL      Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC      CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC      TO BE UBIQUINONE.
CC      -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC      -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC      -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
CC      -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AP050639; AAP05428.1; -
DR      EMBL; AF067166; AAP32450.1; -
DR      EMBL; BC001168; AAH01168.1; -
DR      Genew; HGNC:7697; NDUFB2.
DR      MIM; 603838; -
KM      Oxidoreductase; NAD: ubiquinone: Mitochondrion; Transit peptide.
FT      TRANSIT 1 33 MITOCHONDRION (BY SIMILARITY).
FT      CHAIN 34 105 NADH-UBIQUINONE OXIDOREDUCTASE AGGG
FT      SUBUNIT.
SQ      SEQUENCE 105 AA; 12058 MW; 5698368608CD16F CRC64;

Query Match          10.2%; Score 5; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 PYPPD 15
      11111
Db      85 PYPPD 89

RESULT 45
Y086_CAEEL
ID      Y086_CAEEL STANDARD; PRT; 106 AA.
AC      009238;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      Hypothetical 12.1 kDa protein C18H9.6 in chromosome II.
GN      C18H9.6.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Pelodierinae; Caenorhabditis.
OX      NCBI_Taxid=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Favello T.;
RL      Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: SOME, TO THE C-TERMINAL OF C.ELEGANS T19D12.1.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----

```

```

DR      EMBL; U23147; AAC46690.1; -
DR      WormPep; C18H9.6; CE01806.
KW      Hypothetical protein.
SQ      SEQUENCE 106 AA; 12093 MW; CEC21F8D22BFA863 CRC64;

Query Match          10.2%; Score 5; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 VPYPD 14
      11111
Db      36 VPYPD 40

Search completed: December 4, 2002, 15:49:59
Job time : 12 secs

```





Q925S3  
ID Q925S3 PRELIMINARY; PRT; 147 AA.  
AC Q925S3;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Mus musculus (Mouse).  
DS MRP3.  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RX PubMed=11819679;  
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X., Su C.;  
RT "Mechanism of exogenous nucleic acids and their precursors improving the repair of intestinal epithelium after irradiation in mice.";  
RL World J. Gastroenterol. 6:709-717(2000).  
[2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RA Cui D., Zeng G., Yan X., Li X., Su C.;  
RT "Cloning of mouse genes related to repairing of intestinal epithelium of the irradiated mice by treatment with the intestinal RNA of mice of the same strain.";  
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).  
DR EMBL: AF240166; AAK43731.1; -  
DR InterPro: IPR003006; IG\_MHC.  
DR Pfam: PF00047; Ig; 1.  
SQ SEQUENCE 147 AA; 16274 MW; 800594A12B97191F CRC64;  
Query Match 20.4%; Score 10; DB 11; Length 147;  
Best Local Similarity 100.0%; Pred. No. 0.0025;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAVPPYPDPL 16  
DB 133 GAVPPYPDPL 142  
RESULT 3  
Q925S2  
ID Q925S2 PRELIMINARY; PRT; 170 AA.  
AC Q925S2;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE MRP4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RX PubMed=11819679;  
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X., Su C.;  
RT "Mechanism of exogenous nucleic acids and their precursors improving the repair of intestinal epithelium after irradiation in mice.";  
RL World J. Gastroenterol. 6:709-717(2000).  
[2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RA Cui D., Zeng G., Yan X., Li X., Su C.;  
RT "Cloning of mouse genes related to repairing of intestinal epithelium of the irradiated mice by treatment with the intestinal RNA of mice of the same strain.";  
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).  
DR EMBL: AF240167; AAK43732.1; -  
DR InterPro: IPR003598; Ig\_C2.  
DR

DR InterPro: IPR003006; IG\_MHC.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00408; IGC2; 1.  
KV Immunoglobulin domain.  
SQ SEQUENCE 170 AA; 17978 MW; 5042823CC6C10F38 CRC64;  
Query Match 20.4%; Score 10; DB 11; Length 170;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAVPPYPDPL 16  
DB 156 GAVPPYPDPL 165  
RESULT 4  
Q90YF0  
ID Q90YF0 PRELIMINARY; PRT; 298 AA.  
AC Q90YF0;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CN 8 scfv.  
GN CN 8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C; TISSUE=SPLEEN;  
RX MEDLINE=20183931; PubMed=10706631;  
RA Shinozawa N., Demura T., Fukuda H.;  
RT "Isolation of a vascular cell wall-specific monoclonal antibody recognizing a cell polarity by using a phase display subtraction method.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).  
DR EMBL: AB036341; BAA88633.1; -  
DR HSSP: P01607; IREI.  
DR InterPro: IPR003006; IG\_MHC.  
DR InterPro: IPR003596; IG\_V.  
DR Pfam: PF00047; Ig; 2.  
DR SMART: SM00406; IGV; 2.  
SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;  
Query Match 20.4%; Score 10; DB 11; Length 298;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAVPPYPDPL 16  
DB 284 GAVPPYPDPL 293  
RESULT 5  
Q9XE33  
ID Q9XE33 PRELIMINARY; PRT; 205 AA.  
AC Q9XE33;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Similar to sequence of BAC F7G19 from Arabidopsis thaliana.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzae; Oryza.  
OX NCBI\_TaxID=4350;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsunoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC clone:P0026F07.";  
RT

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP000364; BAA81759.1; -  
 DR InterPro: IPR004822; Histone\_core.  
 DR InterPro: IPR002965; P\_rich\_extensn.  
 DR PRINTS: PR01217; PRICHEXTENS.  
 SQ SEQUENCE 205 AA; 21909 MW; F747D35F886B59C8 CRC64;

Query Match 16.3%; Score 8; DB 10; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 0.55;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVPP 13  
 |||||  
 Db 17 LGAPVPP 24

RESULT 6  
 ID 039303 PRELIMINARY; PRT; 536 AA.  
 AC 039303;  
 DT 01-JUN-1998 (TREMBlrel. 05, Created)  
 DT 01-JUN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Counterpart of HSV-1 gene RL2 and VZV gene 61.  
 GN 63.  
 OS Equine herpesvirus 4.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicellovirus.  
 OX NCBI\_TaxID=10331;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NS80567;  
 RX MEDLINE=98264497; PubMed=9603335.  
 RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;  
 RT "The DNA sequence of equine herpesvirus-4.";  
 RL J. Gen. Virol. 79:1197-1203(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NS80567;  
 RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL: AF030027; AAC59582.1; -  
 DR HSSP: P28990; 1CHC  
 DR InterPro: IPR001841; ZnF\_Ring.  
 DR Pfam: PF00097; zf-C3HC4; 1.  
 DR SMART: SM00184; RING; 1.  
 DR PROSITE: PS00518; ZF\_RING\_1; 1.  
 RL Zinc-finger.  
 SQ SEQUENCE 536 AA; 59686 MW; A973B9B23A92DD08 CRC64;

Query Match 16.3%; Score 8; DB 12; Length 536;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PVPPDPL 16  
 |||||  
 Db 486 PVPPDPL 493

RESULT 7  
 ID 069068 PRELIMINARY; PRT; 268 AA.  
 AC 069068;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE HXI protein (Fragment).  
 GN HXI.  
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=316;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WM8;  
 RX MEDLINE=99008986; PubMed=9791102;  
 RA Metcalf W.W., Wolfe R.S.;  
 RT "Molecular genetic analysis of phosphite and hypophosphite oxidation  
 RT by Pseudomonas stutzeri WM8.";  
 RL J. Bacteriol. 180:5547-5558(1998).  
 CC -1- FUNCTION: BELONGS TO AN OPERON INVOLVED IN HYPOPHOSPHITE  
 CC OXIDATION.  
 CC -1- SIMILARITY: BELONGS TO THE PHN FAMILY.  
 DR EMBL: AF061267; AAC71719.1; -  
 FT NON\_TER 268  
 SQ SEQUENCE 268 AA; 30070 MW; C4ADA5E14C9BDB63 CRC64;

Query Match 14.3%; Score 7; DB 2; Length 268;  
 Best Local Similarity 100.0%; Pred. No. 8.8;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VPYPDPL 16  
 |||||  
 Db 133 VPYPDPL 139

RESULT 8  
 ID 0805K0 PRELIMINARY; PRT; 275 AA.  
 AC 0805K0;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE AGR\_C.880p.  
 GN AGR\_C.880.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gatlung S., Miller N., Blanchard M.,  
 RA Ourullo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Bpp A., Liu F.,  
 RA Wolman C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strud G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT Agrobacterium tumefaciens C58.";  
 RL Science 294:2323-2328(2001).  
 DR EMBL: AE007985; AAK86310.1; -  
 SQ SEQUENCE 275 AA; 29911 MW; CF140634E33C0322 CRC64;

Query Match 14.3%; Score 7; DB 16; Length 275;  
 Best Local Similarity 100.0%; Pred. No. 9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 FOEAYRR 44  
 |||||  
 Db 167 FOEAYRR 173

RESULT 9  
 ID 09HR14 PRELIMINARY; PRT; 286 AA.  
 AC 09HR14;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE 3-hydroxyacyl-CoA dehydrogenase.  
 GN HBD1 OR VNG08816.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
 OC Halobacterium.

OX NCBI\_TaxID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20504483; PubMed=11016950;  
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,  
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
 RA Isebharger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Eberhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
 RT "Genome sequence of Halobacterium species NRC-1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL: AE005014; AAG19174.1; -  
 DR HSSP: P00348; 3HDH.  
 DR InterPro: IPR002135; 3HCDH.  
 DR InterPro: IPR00205; NAD\_binding.  
 DR Pfam: PF00725; 3HCDH\_1.  
 DR Pfam: PF02737; 3HCDH\_N; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 286 AA; 30687 MW; 880F72A7D3AB2342 CRC64;

Query Match 14.3%; Score 7; DB 17; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 9.3;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHIG 37  
 DB 234 ELADHIG 240

RESULT 10  
 OY15C4 PRELIMINARY; PRT; 317 AA.  
 AC OY15C4:  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Probable transcriptional regulator.  
 GN PA0815.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 CC -1- SIMILARITY: TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.  
 DR EMBL: AE004516; AAG04204.1; -  
 DR InterPro: IPR000847; HTH\_LYSR.  
 DR InterPro: IPR005119; LYSR\_subst.  
 DR Pfam: PF00126; HTH\_1; 1.  
 DR Pfam: PF03466; LYSR\_substrate; 1.  
 DR PRINTS: PR00039; HTH\_LYSR.  
 DR PROSITE: PS00044; HTH\_LYSR\_FAMILY; UNKNOWN\_1.  
 KW DNA-binding; Transcription regulation; Complete proteome.  
 SQ SEQUENCE 317 AA; 35990 MW; CC73FDB4A3473960 CRC64;

Query Match 14.3%; Score 7; DB 16; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 OMLGAPV 10  
 DB 42 OMLGAPV 48

RESULT 11  
 OY96H87 PRELIMINARY; PRT; 387 AA.  
 AC OY96H87:  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Unknown (protein for MGC:10474).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL: BC008819; AAH08819.1; -  
 DR InterPro: IPR000923; BlueCu\_1.  
 DR InterPro: IPR000536; Hormone\_rec\_1lg.  
 DR InterPro: IPR000822; Znf\_C2H2.  
 DR InterPro: IPR001628; Znf\_C2steroid.  
 DR Pfam: PF00104; hormone\_rec; 1.  
 DR Pfam: PF00105; zf-C4; 1.  
 DR Prodom: PD000035; Znf\_C2steroid; 1.  
 DR PROSITE: PS00196; COPPER\_BLUE; UNKNOWN\_1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; UNKNOWN\_1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
 KW Zinc-finger.  
 SQ SEQUENCE 387 AA; 43555 MW; 9C96CF2B8E66403C CRC64;

Query Match 14.3%; Score 7; DB 4; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WLGAPVP 11  
 DB 4 WLGAPVP 10

RESULT 12  
 OY18H2 PRELIMINARY; PRT; 467 AA.  
 AC OY18H2:  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE B84.  
 GN B84.  
 OS baboon cytomegalovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=120505;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OCOM4-37;  
 RA Blewett E., Rogers P., Kravitz R., Barry P.;  
 RT "The UL82 gene families of baboon (BacMV) and rhesus cytomegalovirus (rhesus cytomegalovirus) each contain 4 genes and are closely related to the UL82 family (3 genes) of human cytomegalovirus.";  
 RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF411694; AAL10297.1; -  
 DR InterPro: IPR000637; AT\_hook.  
 DR Pfam: PF02178; AT\_hook; 1.  
 SQ SEQUENCE 467 AA; 52651 MW; 335A9A9EFA4831CC CRC64;



Query Match 14.3%; Score 7; DB 12; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYGP 48  
 |||||  
 DB 333 YRRFYGP 339

## RESULT 13

O91CJ5 PRELIMINARY; PRT; 512 AA.  
 AC O91CJ5;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE R84.  
 GN Rhesus cytomegalovirus (strain 68-1) (RhCMV).  
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=103930;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Blewett E.L., Preston R.P., Kravitz R., Barry P.;  
 RA "The UL82 gene families of baboon (baboon cytomegalovirus) and rhesus  
 cytomegalovirus (RhCMV) each contain 4 genes and are closely related  
 to the UL82 family (3 genes) of HCMV."  
 RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-68-1;  
 RA Kravitz R.H., Barry P.A.;  
 RT "Simian cytomegaloviruses as models for HCMV persistence and  
 pathogenesis."  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF078698; AA088223.1; -; EF18BA8354A39656 CRC64;  
 SQ SEQUENCE 512 AA; 57341 MW; EF18BA8354A39656 CRC64;

Query Match 14.3%; Score 7; DB 12; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYGP 48  
 |||||  
 DB 367 YRRFYGP 373

## RESULT 14

O80S15 PRELIMINARY; PRT; 573 AA.  
 AC O80S15;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE U84.  
 OS Chimpazee cytomegalovirus.  
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=188763;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Davison A.J., Akter P., Dolan A., Wright K.M., Addison C.,  
 RA Alencor D.J., Hayward G.S., McGeoch D.J.;  
 RT "The human cytomegalovirus genome revisited."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF480884; AA00723.1; -; B56C2AD35F76D020 CRC64;  
 SQ SEQUENCE 573 AA; 63004 MW; B56C2AD35F76D020 CRC64;

Query Match 14.3%; Score 7; DB 12; Length 573;  
 Best Local Similarity 100.0%; Pred. No. 17;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYGP 48  
 |||||  
 DB 406 YRRFYGP 412

## RESULT 15

O17301 PRELIMINARY; PRT; 4767 AA.  
 AC O17301;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE G01D9.5 protein.  
 GN G01D9.5.  
 OS Caenorhabditis briggsae.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6238;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GUJARAT G16;  
 RA Wu X, Le TT.;  
 RT "The sequence of C. briggsae cosmid G01D9."  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GUJARAT G16;  
 RA Waterston R.;  
 RT "The C. briggsae Genome Sequencing Project."  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U56248; AAA98699.1; -;  
 DR HSSP; P14687; 1AMU.  
 DR InterPro: IPR001227; Ac\_transferase.  
 DR InterPro: IPR000873; AMP-bind.  
 DR InterPro: IPR001242; Condensatn.  
 DR InterPro: IPR000794; ketoacyl-synt.  
 DR InterPro: IPR000734; lipase.  
 DR InterPro: IPR003880; ppanine\_attach.  
 DR InterPro: IPR000379; Ser\_estrs\_site.  
 DR InterPro: IPR000130; zn\_MTpeptidse.  
 DR Pfam: PF00698; Acyl\_transf; 2.  
 DR Pfam: PF00501; AMP-binding; 1.  
 DR Pfam: PF00668; Condensation; 1.  
 DR Pfam: PF00109; ketoacyl-synt; 2.  
 DR Pfam: PF02801; ketoacyl-synt\_C; 2.  
 DR Pfam: PF00550; pp-binding; 5.  
 DR PROSITE: PS50075; ACP\_DOMAIN; 5.  
 DR PROSITE: PS00120; LIPASE\_SER; UNKNOWN\_2.  
 DR PROSITE: PS00012; PHOSPHOPANTHETINE; UNKNOWN\_2.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KW Phosphopantetheine.  
 SQ SEQUENCE 4767 AA; 535178 MW; 0499BB847CB7A07A CRC64;

Query Match 14.3%; Score 7; DB 5; Length 4767;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 DELADHI 36  
 |||||  
 DB 35 DELADHI 41

RESULT 16

O88224 PRELIMINARY; PRT; 56 AA.  
 AC O88224;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Nucleocapsid protein (Fragment).

OS Sin Nombre virus.  
 OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Hantavirus.  
 OX NCBI\_TaxID=37705;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NV-WA-R199;  
 RX MEDLINE=96036486; PubMed-7483255;  
 RA Rowe J.E., St Jeor S.C., Riolo J., Otteson E.W., Monroe M.C.,  
 RA Henderson W.W., Ksiazek T.G., Rollin P.E., Nichol S.T.;  
 RT "Coexistence of several novel hantaviruses in rodents indigenous to  
 North America.";  
 RL Virology 213:122-130(1995).  
 DR EMBL: U33259; AAC36797.1; -;  
 DR InterPro: IPR002214; Hanta\_nucleocap.  
 DR Pfam: PF00846; Hanta\_nucleocap; 1.  
 DR Prodom: PD001501; Hanta\_nucleocap; 1.  
 FT NON\_TER 1 1  
 FT SEQUENCE 56 AA; 6107 MW; 15BE701F402F7D3B CRC64;

Query Match 12.2%; Score 6; DB 12; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
 |||||  
 DB 23 ELADHI 28

## RESULT 17

OQ82160 PRELIMINARY; PRT; 56 AA.  
 AC Q82160; O82161;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Nucleocapsid protein (Fragment).  
 OS unidentified.  
 OC unidentified.  
 OX NCBI\_TaxID=32644;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SIN-NOMBRE/CA-MO-R159, AND SIN-NOMBRE/NV-NY-R301;  
 RX MEDLINE=96130200; PubMed-8553562;  
 RA Henderson W.W., Monroe M.C., St Jeor S.C., Thayer W.P., Rowe J.E.,  
 RA Peters C.J., Nichol S.T.;  
 RT "Naturally occurring Sin Nombre virus genetic reassortants.";  
 RL Virology 214:602-610(1995).  
 DR EMBL: U45015; AAB48145.1; -;  
 DR EMBL: U45016; AAB48146.1; -;  
 DR InterPro: IPR002214; Hanta\_nucleocap.  
 DR Pfam: PF00846; Hanta\_nucleocap; 1.  
 DR Prodom: PD001501; Hanta\_nucleocap; 1.  
 KW Nucleocapsid.  
 FT NON\_TER 1 1  
 FT SEQUENCE 56 AA; 6107 MW; 15BE701F402F7D3B CRC64;

Query Match 12.2%; Score 6; DB 12; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
 |||||  
 DB 23 ELADHI 28

## RESULT 18

OQ8R2M7 PRELIMINARY; PRT; 99 AA.  
 AC O8R2M7;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE B1168H06.7 protein.  
 GN B1168H06.7.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC  
 clone:B1168H06.";  
 RL submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003563; BAB89169.1; -;  
 SO SEQUENCE 99 AA; 10491 MW; 257199E2438B5E6 CRC64;

Query Match 12.2%; Score 6; DB 10; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
 |||||  
 DB 25 LGAPVP 30

## RESULT 19

OQ27316 PRELIMINARY; PRT; 100 AA.  
 AC Q27316;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE L71-1.  
 GN EIG71EA OR L71-1 OR CG16931.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OREGON-R; TISSUE=SALIVARY GLAND;  
 RA Wright L.G., Chen T., Thummel C.S., Guild G.M.;  
 RL Submitted (Apr-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON-S;  
 RA Wright L.G., Chen T., Thummel C.S., Guild G.M.;  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U24095; AAA65109.1; -;  
 DR EMBL: U23836; AAA74176.1; -;  
 DR Flybase: FBgn000458; EIG71EA.  
 DR InterPro: IPR003475; Insect\_unk.  
 DR Pfam: PF02448; L71; 1.  
 SO SEQUENCE 100 AA; 11895 MW; 0D7A22A0639D38A8 CRC64;

Query Match 12.2%; Score 6; DB 5; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 DCDELA 33  
 |||||  
 DB 24 DCDELA 29

## RESULT 20

OQ9VUS3 PRELIMINARY; PRT; 100 AA.  
 AC Q9VUS3;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

```

DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)
DE EIG71EA protein.
GN EIG71EA OR CG16931.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephyroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abri J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Fosler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacile J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003530; AAF49602.1; -.
DR FlyBase: FBgn0004588; EIG71EA.
DR InterPro: IPR003475; insect_Unk.
DR Pfam: PF02448; L71; 1.
SQ SEQUENCE 100 AA; 11867 MW; 1B1D34D1048B5FA8 CRC64;

Query Match 12.2%; Score 6; DB 5; Length 100;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 DCDELA 33
DB 24 DCDELA 29

RESULT 21
Q41032 PRELIMINARY; PRT; 136 AA.
AC Q41032:
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE A550R protein.

```

```

GN A550R.
OS Parametium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98022962; PubMed=9356347;
RA Li Y., Lu Z., Sun L., Ropp S., Kutish G.F., Rock D.L., Van Etten J.L.;
RT "Analysis of 74 kb of DNA located at the right end of the 330-kb
RT chlorella virus PBCV-1 genome.";
RL Virology 237:360-377(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20033326; PubMed=10544099;
RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
RA Lisek A.D., Nickerson K.W., Van Etten J.L.;
RT "Chlorella virus PBCV-1 encodes a functional homospemidine
RT synthase.";
RL Virology 263:254-262(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20478054; PubMed=11021991;
RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
RT PBCV-1.";
RL Virology 276:27-36(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RA Gurnon J.R., Graves M.V., Van Etten J.L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U42580; AAC96997.1; -.
SQ SEQUENCE 136 AA; 15566 MW; 18AFC2D446040481 CRC64;

Query Match 12.2%; Score 6; DB 12; Length 136;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
DB 12 LGAPVP 17

RESULT 22
Q41658 PRELIMINARY; PRT; 143 AA.
AC Q41658:
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE NADH ubiquinone oxidoreductase subunit 5 (nad5).
OS Vicia faba (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

OC Spermatothya; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eukaryotes; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.  
 OX NCBI\_TaxID=3906;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AD23/N FERTILE LINE;  
 RA Scheepers D.G.J.M., Luo H., Boutry M.;  
 RT "Variant mitochondrial transcripts of a broad bean line are associated  
 with two point mutations located upstream of nad5 exon c.";  
 RL Plant Sci. 129:203-212(1997).  
 DR EMBL: L36945; AAB97304.1; -;  
 KW Mitochondrion; Ubiquinone.  
 SQ SEQUENCE 143 AA; 15455 MW; 577847CA88C0DFB4 CRC64;  
 Query Match 12.2%; Score 6; DB 10; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 8 APVYP 13  
 DB 55 APVYP 60  
 RESULT 23  
 O87779 PRELIMINARY; PRT; 166 AA.  
 AC O87779;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical 18.1 kDa protein (Fragment).  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19698;  
 RX MEDLINE=93328703; PubMed=8335649;  
 RA Gijlot P., De Kessel M., Coene M., Machtelinckx L., Cocito C.;  
 RT "Isolation and sequencing of the gene coding for an antigenic 34-  
 kDa protein of Mycobacterium paratuberculosis.";  
 RL J. Bacteriol. 175:4930-4935(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19698;  
 RX MEDLINE=99159670; PubMed=10068253;  
 RA Gijlot P.;  
 RT "Specificity of the 34-kDa protein immunodominant protein of  
 Mycobacterium avium subsp. paratuberculosis.";  
 RL Clin. Diagn. Lab. Immunol. 6:146-148(1999).  
 DR EMBL: X68102; CAA48222.1; -;  
 DR InterPro: IPR002103; BacLuciferase.  
 DR Pfam: PF00296; bacLuciferase; 1.  
 KW Hypothetical protein.  
 FT NON\_TER 166  
 SQ SEQUENCE 166 AA; 18052 MW; B7A8F0F7A6A55344 CRC64;  
 Query Match 12.2%; Score 6; DB 2; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 30 DELADH 35  
 DB 123 DELADH 128  
 RESULT 24  
 O47896 PRELIMINARY; PRT; 169 AA.  
 AC O47896;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ORF 169.  
 OS Tolypothrix sp. PCC 7601.  
 OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Tolypothrix.  
 OX NCBI\_TaxID=1188;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=88260883; PubMed=2838727;  
 RA Parsot C., Mazel D.;  
 RT "Cloning and nucleotide sequence of the thrB gene from the  
 cyanobacterium Calothrix PCC 7601.";  
 RL Mol. Microbiol. 1:45-52(1987).  
 DR EMBL: Y00522; CAA68577.1; -;  
 SQ SEQUENCE 169 AA; 19182 MW; 1A036CFA6CEA4A00 CRC64;  
 Query Match 12.2%; Score 6; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 40 EAYRF 45  
 DB 43 EAYRF 48  
 RESULT 25  
 O8ZJ02 PRELIMINARY; PRT; 185 AA.  
 AC O8ZJ02;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Putative fibriol subunit.  
 GN STD OR STM4592.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Salmonella.  
 OX NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porrolik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvanev E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of salmonella enterica serovar Typhimurium  
 LT2.";  
 RL Nature 413:852-856(2001).  
 DR EMBL: AE008916; AAL23407.1; -;  
 DR InterPro: IPR000259; Fimbrin.  
 DR Pfam: PF00419; Fimbrin; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 185 AA; 19337 MW; C83C8701EBA9676E CRC64;  
 Query Match 12.2%; Score 6; DB 16; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 10 VYPDP 15  
 DB 99 VYPDP 104  
 RESULT 26  
 O8Z0S9 PRELIMINARY; PRT; 185 AA.  
 AC O8Z0S9;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Putative fibriol subunit.  
 GN STD OR STM4940.

```

OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Salmonella.
RN NCBI_Taxid=601;
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque N., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Goara P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.; a multiple drug resistant Salmonella
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL: AL627284; CAD03423.1;
DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 185 AA; 19425 MW; 836A210DEA432234 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 185;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 VPPDP 15
DB 99 VPPDP 104

RESULT 27
O9FA19 PRELIMINARY; PRT; 198 AA.
ID O9FA19;
AC O9FA19;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Putative transcription regulator.
GN TETR.
OS Mycobacterium paratuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_Taxid=1770;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20431891; PubMed=10974106;
RA Bull T.J., Hermion-Taylor J., Pavlik I., El-Zaatari F., Tizard M.;
RT "Characterization of IS900 loci in Mycobacterium avium subsp.
RT paratuberculosis and development of multiplex PCR typing."
RL Microbiology 146:2185-2197(2000).
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL: AJ250023; CAC10267.1;
DR InterPro: IPR001647; HTH_Tetr.
DR Pfam: PF00440; tetr. 1.
DR PRINTS: PR00455; HTHTETR.
KM DNA-binding; Transcription regulation.
SQ SEQUENCE 198 AA; 21553 MW; 06786AEE165F302 CRC64;

Query Match 12.2%; Score 6; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 OEAYR 44
DB 144 OEAYR 149

RESULT 28

```

```

O8YN80 PRELIMINARY; PRT; 214 AA.
ID O8YN80;
AC O8YN80;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein Al14690.
GN Al14690.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_Taxid=103650;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2159285; PubMed=11759840;
RA Kaneke T., Nakamura Y., Nolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Itiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL: AP003597; BAB/6389.1;
DR InterPro: IPR001601; Methyltransf.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 214 AA; 24688 MW; 7A4D1DC45E40F947 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 OMLGP 9
DB 196 OMLGP 201

RESULT 29
O8Y1H6 PRELIMINARY; PRT; 217 AA.
ID O8Y1H6;
AC O8Y1H6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Probable riboflavin synthase (Alpha chain) protein (EC 2.5.1.9).
GN RIBE OR RSC0714 OR RS05139.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_Taxid=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM1100;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mangenot S.,
RA Ariat M., Billault A., Brotier P., Camus J.C., Catalicio L.,
RA Chandler M., Chisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguler P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL: AL646060; CAD14244.1;
DR InterPro: IPR001783; Lum_binding.
DR Pfam: PF00677; Lum_binding. 2.
DR PRODOM: PD004110; Lum_binding. 1.
DR TIGRFAMs: TIGR00187; ribe. 1.
KM Transferrase; Complete proteome.
SQ SEQUENCE 217 AA; 22663 MW; 4468296091583172 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 217;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      32 LADHIG 37
      |||||
Db      92 LADHIG 97

RESULT 30
O8Y1B2 PRELIMINARY; PRT; 221 AA.
AC O8Y1B2:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Hypothetical protein RSC0778.
GN RSC0778 OR RS05076.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brothier P., Camus J.C., Cattolico L.,
RA Chandelier M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Laye M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Trebault P., Whalen M., Winkler P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
KW EMBL: AL646061; CAD14480.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 221 AA; 2486 MW; A9EB517D9402E246 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 221;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 LGAPVP 11
      |||||
Db      32 LGAPVP 37

RESULT 31
O07473 PRELIMINARY; PRT; 231 AA.
AC O07473:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE GdmF.
GN GDMF.
OS Staphylococcus gallinarum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1293;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TUE3928;
RA MEDLINE=97294510; PubMed=9150266;
RA Schnell N., Entian K.D., Gotz F., Horner T., Kellner R., Jung G.;
RT "Structural gene isolation and prepeptide sequence of gallidermin, a
RT new lantibiotic containing antibiotic.";
RL FEMS Microbiol. Lett. 49:263-267(1989).
RN [2].
RP SEQUENCE FROM N.A.
RX STRAIN=TUE3928;
RA MEDLINE=97294510; PubMed=9150266;
RA Peschel A., Schnell N., Hille M., Entian K.D., Gotz F.;
RT "Secretion of the lantibiotics epidermin and gallidermin: sequence
RT analysis of the genes gdmT and gdmH, their influence on epidermin
RT production and their regulation by EpiO.";
RL Mol. Gen. Genet. 254:312-318(1997).

```

```

DR EMBL: U61158; AAB61132.1; -.
DR InterPro: IPR003593; AAA_Arase.
DR InterPro: IPR003439; ABC_transportr.
DR Pfam: PF00005; ABC_tran.1.
DR ProDom: PD000006; ABC_transportr; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW ATP-binding.
SQ SEQUENCE 231 AA; 25902 MW; 2475D9B3F5D237A0 CRC64;

Query Match 12.2%; Score 6; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      32 LADHIG 37
      |||||
Db      195 LADHIG 200

RESULT 32
O93HB2 PRELIMINARY; PRT; 234 AA.
AC O93HB2:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE ClpX homolog.
GN ClpX.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: Deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL: AB070947; BAB69277.1; -.
DR InterPro: IPR004176; Clp_N.
DR Pfam: PF02861; Clp_N; 2.
SQ SEQUENCE 234 AA; 24300 MW; AFC1EF7EB85C4B7C CRC64;

Query Match 12.2%; Score 6; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      32 LADHIG 37
      |||||
Db      34 LADHIG 39

RESULT 33
O9K0X9 PRELIMINARY; PRT; 235 AA.
AC O9K0X9:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Cell division protein FtsQ.
GN NMB0425.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MC58 / SEROGROUP B;
RA MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,

```

RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
 RA Nelson W.C., Gwin M.L., Deboy R., Peterson J.D., Hickey E.K.,  
 RA Haft D.H., Salberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
 RA Mason T., Cleeke A., Parisey D.S., Blair E., Ciltone H., Clark E.B.,  
 RA Cotton M.D., Uteback T.R., Khouri H., Qin H., Vamathevan J.,  
 RA Gill J., Scarlato V., Maignani V., Pilza M., Grandi G., Sun L.,  
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappaport R., Venter J.C.;  
 RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain  
 MC58.";  
 RT Science 287:1809-1815(2000).  
 DR EMBL: AEO02398; AAF40863.1; -.  
 DR TIGR: NMB0425; -.  
 KW Complete proteome.  
 SO SEQUENCE 235 AA; 27135 MW; B7E605926BD5E9FF CRC64;

Query Match 12.2%; Score 6; DB 16; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 QEAYRR 44  
 Db 70 QEAYRR 75

RESULT 34  
 O9JUT00 PRELIMINARY; PRT; 242 AA.

AC O9JUT00;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Cell division protein.  
 GN FTSQ OR NMA2059.  
 OS *Neisseria meningitidis* (serogroup A).  
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=65699;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=42491 / SEROGROUP A / SEROTYPE 4A;  
 RX MEDLINE=20222556; PubMed=10761919;  
 RA Parhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,  
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
 RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holroyd S.,  
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,  
 RA Rajandream M.A., Rutherford K.M., Simmonds K., Skelton J.,  
 RA Whitehead S., Spratt B.G., Barrall B.G.;  
 RT "Complete DNA sequence of a serogroup A strain of *Neisseria*  
 meningitidis 22491.";  
 RL Nature 404:502-506(2000).  
 DR EMBL: AL162758; CAB85277.1; -.  
 KW Complete proteome.  
 SO SEQUENCE 242 AA; 27933 MW; A29018736D2EFB35 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 QEAYRR 44  
 Db 77 QEAYRR 82

RESULT 35  
 O9VXM4 PRELIMINARY; PRT; 248 AA.

AC O9VXM4;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE CG9066 protein (LID2946P).  
 GN CG9066.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman J.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideygam C.,  
 RA Jaimali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,  
 RA Kamel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AEO03500; AAP48534.1; -.  
 DR EMBL: AY061163; AA128711.1; -.  
 DR FLYbase: FBgn0030703; CG9066.  
 DR InterPro: IPR001199; Cyt-B5.  
 DR Pfam: PF001173; heme\_1; 1.  
 SO SEQUENCE 248 AA; 27921 MW; 1A7B9C67BDE72FA3 CRC64;

Query Match 12.2%; Score 6; DB 5; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RREYGP 48  
 Db 115 RREYGP 120

RESULT 36  
 O8YJF8 PRELIMINARY; PRT; 252 AA.

AC 0810F8:  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Acetyltransferase (EC 2.3.1.-)  
 GN BME0125.  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29459;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA DeIvecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujer C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,  
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Haeckl R., Kypides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
 DR EMBL: AE009455; AAL51307.1; -  
 DR InterPro: IPR000182; GCM5acetyltransf.  
 DR Pfam: PF00583; Acetyltransf.1.  
 KW Transferase; Acyltransferase; Complete proteome.  
 SQ SEQUENCE 252 AA; 27861 MW; 5A9582613939E773 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 252;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 37 GFEAY 42  
 |||||  
 DB 237 GFEAY 242

RESULT 37  
 088025 PRELIMINARY; PRT; 264 AA.  
 AC 088025;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative transcriptional regulator.  
 GN SCO669 OR SC5A7.19C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Crocin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -1- SIMILARITY: BELONGS TO THE ICLR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 DR EMBL: AL031107; CAA19948.1; -  
 DR InterPro: IPR00285; HTH\_ICLR.  
 DR Pfam: PF01614; ICLR.1.  
 DR SMART: SMO0346; HTH\_ICLR.1.  
 KW DNA-binding; Transcription regulation.  
 SQ SEQUENCE 264 AA; 28066 MW; 99C0F97B015D3709 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 FDEAYR 43  
 |||||  
 DB 153 FDEAYR 158

RESULT 38  
 08TPA8 PRELIMINARY; PRT; 273 AA.  
 AC 08TPA8;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical protein MA2005.  
 GN MA2005.  
 OS Methanosarcina acetivorans.  
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;  
 OC Methanosarcinaceae; Methanosarcina.  
 OX NCBI\_TaxID=2214;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
 RX MEDLINE=21929760; PubMed=11932238;  
 RA Galagan J.E., Nusbaum C., Roy A., Smirnov S., Alnoor D., Brown A.,  
 RA Fitzhugh W., Calvo S., Engels R., Endrizzi M.G., Macdonald P.,  
 RA Allen N., Naylor J., Stange-Thomann N., DeAngelis K., Johnson R.,  
 RA Linton L., McEwan P., McKernan K., Talamas J., Turrell A., Ye W.,  
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,  
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,  
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,  
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
 RA Metcalf W.W., Birren B.;  
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
 RT and physiological diversity.";  
 RL Genome Res. 12:532-542(2002).  
 DR EMBL: AE010865; AAM05408.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 273 AA; 31581 MW; 8C2058B9831AA22A CRC64;

Query Match 12.2%; Score 6; DB 17; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRFTG 47  
 |||||  
 DB 266 YRFTG 271



```

RESULT 39
Q92MP0
ID 092MP0 PRELIMINARY; PRT; 287 AA.
AC 092MP0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE Putative transport system permease ABC transporter protein.
GN R02573 OR SMC02345.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL: AF591791; CAC47152.1; -.
DR InterPro: IPR000515; BPD_transp.
DR Pfam: PF00528; BPD_transp. 1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 31044 MW; BD1E31DEBA93EE7A CRC64;

Query Match 12.2%; Score 6; DB 16; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 LADHIG 37
Db 84 LADHIG 89

RESULT 40
Q9A3K1
ID Q9A3K1 PRELIMINARY; PRT; 289 AA.
AC Q9A3K1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE Hypothetical protein CC3203.
GN CC3203.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nielsen W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eiken J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri J., Shetty J., Berry K.,
RA Ullrichback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005984; AAK25165.1; -.
DR TIGR: CC3203; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 289 AA; 31055 MW; 8DB2411C8AB7CFA7 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 289;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 QEAYRR 44
Db 214 QEAYRR 219

RESULT 41
Q27855
ID Q27855 PRELIMINARY; PRT; 304 AA.
AC Q27855;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE Quinolinate synthetase.
GN MTH1827.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-T., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL: AE000936; AAB86293.1; -.
DR InterPro: IPR003473; NADA.
DR Pfam: PF02445; NADA; 1.
DR TIGRPFAMs: TIGR00550; nada; 1.
KW Complete proteome.
SQ SEQUENCE 304 AA; 34393 MW; B5C48ACE482143DD CRC64;

Query Match 12.2%; Score 6; DB 17; Length 304;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 ELADHI 36
Db 208 ELADHI 213

RESULT 42
P74835
ID P74835 PRELIMINARY; PRT; 318 AA.
AC P74835;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, last annotation update)
DE Hypothetical 34.1 kDa protein.
OS Sphingomonas sp. 588.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC Sphingomonas.
OX NCBI_TaxID=46624;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=588;
RX MEDLINE=96196177; PubMed=8626338;
RA Yamazaki M., Thorne L., Mikolajczak M., Armentrout R.W., Pollock T.J.;
RT "Linkage of genes essential for synthesis of a polysaccharide capsule
RT in Sphingomonas strain 588."
RL J. Bacteriol. 178:2676-2687(1996).
DR EMBL: U51197; AAC44077.1; -.
DR InterPro: IPR000620; DUF6.
DR Pfam: PF00892; DUF6; 2.

```

KW Hypothetical protein.  
SQ SEQUENCE 318 AA; 34143 MW; A503E62A8BD0363C CRC64;

Query Match 12.2%; Score 6; DB 2; Length 318;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WIGAPV 10  
|||||  
Db 286 WIGAPV 291

## RESULT 43

ID Q9SLX6 PRELIMINARY; PRT; 318 AA.

AC Q9SLX6;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE TRABL.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
OS [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=CV. NIPPONBARE;  
RX MEDLINE=20079656; PubMed=10611387;  
RA Hobo T., Koyama Y., Hattori T.;  
RT "A bz1p factor, TRABL, interacts with Vp1 and mediates abscisic acid-  
induced transcription.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:15348-15353(1999).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY.  
DR EMBL: AB023288; BAA83740.1; -  
DR TRANSFAC: T04821; -  
DR InterPro: IPR004827; TF\_BZIP.  
DR Pfam: PF00170; bZIP; 1.  
DR SMART: SM00338; BRLZ; 1.  
DR PROSITE: PS00036; BZIP\_BASIC; 1.  
KW DNA-binding; Nuclear protein.  
SQ SEQUENCE 318 AA; 33761 MW; 76935F70B3AD74A5 CRC64;

Query Match 12.2%; Score 6; DB 10; Length 318;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APVYP 13  
|||||  
Db 207 APVYP 212

## RESULT 44

ID Q8SY52 PRELIMINARY; PRT; 329 AA.

AC Q8SY52;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE GH13458p.  
GN ELA.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
OS [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=BERKELEY;  
RA Stapleton M., Brokstein P., Hong L., Agdayani A., Carlson J.,  
Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanenvong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY075334; AAL68200.1; -  
SQ SEQUENCE 329 AA; 33199 MW; 62E9C1B7EC09302 CRC64;

Query Match 12.2%; Score 6; DB 5; Length 329;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APVYP 13  
|||||  
Db 135 APVYP 140

## RESULT 45

ID Q9RM60 PRELIMINARY; PRT; 332 AA.

AC Q9RM60;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical protein DR0809.  
GN DR0809.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/deinococcus group; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=1293;  
OS [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=RL;  
RC MEDLINE=20036896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Halt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
radiodurans RL.";  
RL Science 286:1571-1577(1999).  
DR EMBL: AF001935; AAF10387.1; -  
DR TIGR: DR0809;  
DR InterPro: IPR000537; UDA.  
DR Pfam: PF01040; UDA; 1.  
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 332 AA; 35374 MW; 01EC6E2ADC10E220 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 332;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPV 11  
|||||  
Db 204 LGAPV 209

Search completed: December 4, 2002, 15:50:56  
Job time : 29 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:49:46 ; Search time 141 Seconds

(Without alignments)

224.056 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 49

Sequence: 1 YLQWIGAPVPYPPDPLKPRR.....DELADHIGFQEAATRRFGVPV 49

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 4569144 seqs, 644733110 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending\_Patents\_MA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCRTUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep:\*  
9: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*  
10: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep:\*  
11: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep:\*  
12: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep:\*  
13: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep:\*  
14: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep:\*  
15: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep:\*  
16: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep:\*  
17: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep:\*  
18: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep:\*  
19: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep:\*  
20: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep:\*  
21: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep:\*  
22: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep:\*  
23: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep:\*  
24: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep:\*  
25: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep:\*  
26: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep:\*  
27: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	93.9	49	13	US-08-973-667-4
2	25	51.0	42	14	US-09-036-085-5
3	25	51.0	46	20	US-09-657-276-1095
4	25	51.0	47	20	US-09-657-276-1093
5	25	51.0	49	1	PCT-US02-22821-108
6	25	51.0	49	13	US-08-973-667-1

7	25	51.0	49	13	US-08-973-667-2	Sequence 2, App11
8 <td>25<td>51.0<td>49<td>13<td>US-08-973-667-3</td><td>Sequence 3, App11</td></td></td></td></td>	25 <td>51.0<td>49<td>13<td>US-08-973-667-3</td><td>Sequence 3, App11</td></td></td></td>	51.0 <td>49<td>13<td>US-08-973-667-3</td><td>Sequence 3, App11</td></td></td>	49 <td>13<td>US-08-973-667-3</td><td>Sequence 3, App11</td></td>	13 <td>US-08-973-667-3</td> <td>Sequence 3, App11</td>	US-08-973-667-3	Sequence 3, App11
9 <td>25<td>51.0<td>49<td>18<td>US-09-462-931-2</td><td>Sequence 2, App11</td></td></td></td></td>	25 <td>51.0<td>49<td>18<td>US-09-462-931-2</td><td>Sequence 2, App11</td></td></td></td>	51.0 <td>49<td>18<td>US-09-462-931-2</td><td>Sequence 2, App11</td></td></td>	49 <td>18<td>US-09-462-931-2</td><td>Sequence 2, App11</td></td>	18 <td>US-09-462-931-2</td> <td>Sequence 2, App11</td>	US-09-462-931-2	Sequence 2, App11
10 <td>25<td>51.0<td>49<td>25<td>US-10-197-954-108</td><td>Sequence 108, App</td></td></td></td></td>	25 <td>51.0<td>49<td>25<td>US-10-197-954-108</td><td>Sequence 108, App</td></td></td></td>	51.0 <td>49<td>25<td>US-10-197-954-108</td><td>Sequence 108, App</td></td></td>	49 <td>25<td>US-10-197-954-108</td><td>Sequence 108, App</td></td>	25 <td>US-10-197-954-108</td> <td>Sequence 108, App</td>	US-10-197-954-108	Sequence 108, App
11 <td>25<td>51.0<td>98<td>3</td><td>US-07-717-811A-5</td><td>Sequence 5, App11</td></td></td></td>	25 <td>51.0<td>98<td>3</td><td>US-07-717-811A-5</td><td>Sequence 5, App11</td></td></td>	51.0 <td>98<td>3</td><td>US-07-717-811A-5</td><td>Sequence 5, App11</td></td>	98 <td>3</td> <td>US-07-717-811A-5</td> <td>Sequence 5, App11</td>	3	US-07-717-811A-5	Sequence 5, App11
12 <td>25<td>51.0<td>98<td>6</td><td>US-08-246-626-5</td><td>Sequence 5, App11</td></td></td></td>	25 <td>51.0<td>98<td>6</td><td>US-08-246-626-5</td><td>Sequence 5, App11</td></td></td>	51.0 <td>98<td>6</td><td>US-08-246-626-5</td><td>Sequence 5, App11</td></td>	98 <td>6</td> <td>US-08-246-626-5</td> <td>Sequence 5, App11</td>	6	US-08-246-626-5	Sequence 5, App11
13 <td>25<td>51.0<td>100<td>1</td><td>PCT-US01-08655-186</td><td>Sequence 186, App</td></td></td></td>	25 <td>51.0<td>100<td>1</td><td>PCT-US01-08655-186</td><td>Sequence 186, App</td></td></td>	51.0 <td>100<td>1</td><td>PCT-US01-08655-186</td><td>Sequence 186, App</td></td>	100 <td>1</td> <td>PCT-US01-08655-186</td> <td>Sequence 186, App</td>	1	PCT-US01-08655-186	Sequence 186, App
14 <td>25<td>51.0<td>100<td>1</td><td>PCT-US01-12010-3</td><td>Sequence 3, App11</td></td></td></td>	25 <td>51.0<td>100<td>1</td><td>PCT-US01-12010-3</td><td>Sequence 3, App11</td></td></td>	51.0 <td>100<td>1</td><td>PCT-US01-12010-3</td><td>Sequence 3, App11</td></td>	100 <td>1</td> <td>PCT-US01-12010-3</td> <td>Sequence 3, App11</td>	1	PCT-US01-12010-3	Sequence 3, App11
15 <td>25<td>51.0<td>127<td>25<td>US-10-143-899-23</td><td>Sequence 23, App1</td></td></td></td></td>	25 <td>51.0<td>127<td>25<td>US-10-143-899-23</td><td>Sequence 23, App1</td></td></td></td>	51.0 <td>127<td>25<td>US-10-143-899-23</td><td>Sequence 23, App1</td></td></td>	127 <td>25<td>US-10-143-899-23</td><td>Sequence 23, App1</td></td>	25 <td>US-10-143-899-23</td> <td>Sequence 23, App1</td>	US-10-143-899-23	Sequence 23, App1
16 <td>25<td>51.0<td>127<td>26<td>US-10-217-651-336</td><td>Sequence 336, App</td></td></td></td></td>	25 <td>51.0<td>127<td>26<td>US-10-217-651-336</td><td>Sequence 336, App</td></td></td></td>	51.0 <td>127<td>26<td>US-10-217-651-336</td><td>Sequence 336, App</td></td></td>	127 <td>26<td>US-10-217-651-336</td><td>Sequence 336, App</td></td>	26 <td>US-10-217-651-336</td> <td>Sequence 336, App</td>	US-10-217-651-336	Sequence 336, App
17 <td>23<td>46.9<td>42<td>27<td>US-60-160-203-3372</td><td>Sequence 3372, App</td></td></td></td></td>	23 <td>46.9<td>42<td>27<td>US-60-160-203-3372</td><td>Sequence 3372, App</td></td></td></td>	46.9 <td>42<td>27<td>US-60-160-203-3372</td><td>Sequence 3372, App</td></td></td>	42 <td>27<td>US-60-160-203-3372</td><td>Sequence 3372, App</td></td>	27 <td>US-60-160-203-3372</td> <td>Sequence 3372, App</td>	US-60-160-203-3372	Sequence 3372, App
18 <td>23<td>46.9<td>42<td>27<td>US-60-163-123-1219</td><td>Sequence 1219, App</td></td></td></td></td>	23 <td>46.9<td>42<td>27<td>US-60-163-123-1219</td><td>Sequence 1219, App</td></td></td></td>	46.9 <td>42<td>27<td>US-60-163-123-1219</td><td>Sequence 1219, App</td></td></td>	42 <td>27<td>US-60-163-123-1219</td><td>Sequence 1219, App</td></td>	27 <td>US-60-163-123-1219</td> <td>Sequence 1219, App</td>	US-60-163-123-1219	Sequence 1219, App
19 <td>23<td>46.9<td>42<td>27<td>US-60-169-840-5089</td><td>Sequence 5089, App</td></td></td></td></td>	23 <td>46.9<td>42<td>27<td>US-60-169-840-5089</td><td>Sequence 5089, App</td></td></td></td>	46.9 <td>42<td>27<td>US-60-169-840-5089</td><td>Sequence 5089, App</td></td></td>	42 <td>27<td>US-60-169-840-5089</td><td>Sequence 5089, App</td></td>	27 <td>US-60-169-840-5089</td> <td>Sequence 5089, App</td>	US-60-169-840-5089	Sequence 5089, App
20 <td>19<td>38.8</td><td>140</td><td>1</td><td>PCT-US01-08655-478</td><td>Sequence 478, App</td></td>	19 <td>38.8</td> <td>140</td> <td>1</td> <td>PCT-US01-08655-478</td> <td>Sequence 478, App</td>	38.8	140	1	PCT-US01-08655-478	Sequence 478, App
21 <td>16<td>32.7</td><td>21</td><td>3</td><td>US-07-717-811A-10</td><td>Sequence 10, App1</td></td>	16 <td>32.7</td> <td>21</td> <td>3</td> <td>US-07-717-811A-10</td> <td>Sequence 10, App1</td>	32.7	21	3	US-07-717-811A-10	Sequence 10, App1
22 <td>16<td>32.7</td><td>21</td><td>6</td><td>US-08-246-626-10</td><td>Sequence 10, App1</td></td>	16 <td>32.7</td> <td>21</td> <td>6</td> <td>US-08-246-626-10</td> <td>Sequence 10, App1</td>	32.7	21	6	US-08-246-626-10	Sequence 10, App1
23 <td>14<td>28.6</td><td>15</td><td>3</td><td>US-07-717-811A-12</td><td>Sequence 12, App1</td></td>	14 <td>28.6</td> <td>15</td> <td>3</td> <td>US-07-717-811A-12</td> <td>Sequence 12, App1</td>	28.6	15	3	US-07-717-811A-12	Sequence 12, App1
24 <td>14<td>28.6</td><td>15</td><td>6</td><td>US-08-246-626-12</td><td>Sequence 12, App1</td></td>	14 <td>28.6</td> <td>15</td> <td>6</td> <td>US-08-246-626-12</td> <td>Sequence 12, App1</td>	28.6	15	6	US-08-246-626-12	Sequence 12, App1
25 <td>13<td>26.5</td><td>13</td><td>17</td><td>US-09-341-590-44</td><td>Sequence 44, App1</td></td>	13 <td>26.5</td> <td>13</td> <td>17</td> <td>US-09-341-590-44</td> <td>Sequence 44, App1</td>	26.5	13	17	US-09-341-590-44	Sequence 44, App1
26 <td>13<td>26.5</td><td>13</td><td>20</td><td>US-09-657-276-1096</td><td>Sequence 1096, App</td></td>	13 <td>26.5</td> <td>13</td> <td>20</td> <td>US-09-657-276-1096</td> <td>Sequence 1096, App</td>	26.5	13	20	US-09-657-276-1096	Sequence 1096, App
27 <td>11<td>22.4</td><td>13</td><td>13</td><td>US-08-973-667-6</td><td>Sequence 6, App11</td></td>	11 <td>22.4</td> <td>13</td> <td>13</td> <td>US-08-973-667-6</td> <td>Sequence 6, App11</td>	22.4	13	13	US-08-973-667-6	Sequence 6, App11
28 <td>11<td>22.4</td><td>43</td><td>14</td><td>US-09-036-085-22</td><td>Sequence 22, App1</td></td>	11 <td>22.4</td> <td>43</td> <td>14</td> <td>US-09-036-085-22</td> <td>Sequence 22, App1</td>	22.4	43	14	US-09-036-085-22	Sequence 22, App1
29 <td>11<td>22.4</td><td>74</td><td>1</td><td>PCT-US02-30312-2134</td><td>Sequence 2134, App</td></td>	11 <td>22.4</td> <td>74</td> <td>1</td> <td>PCT-US02-30312-2134</td> <td>Sequence 2134, App</td>	22.4	74	1	PCT-US02-30312-2134	Sequence 2134, App
30 <td>11<td>22.4</td><td>74</td><td>1</td><td>PCT-US02-30312-2134</td><td>Sequence 2134, App</td></td>	11 <td>22.4</td> <td>74</td> <td>1</td> <td>PCT-US02-30312-2134</td> <td>Sequence 2134, App</td>	22.4	74	1	PCT-US02-30312-2134	Sequence 2134, App
31 <td>11<td>22.4</td><td>74</td><td>23</td><td>US-09-962-756-2134</td><td>Sequence 2134, App</td></td>	11 <td>22.4</td> <td>74</td> <td>23</td> <td>US-09-962-756-2134</td> <td>Sequence 2134, App</td>	22.4	74	23	US-09-962-756-2134	Sequence 2134, App
32 <td>11<td>22.4</td><td>74</td><td>26</td><td>US-10-253-471-2134</td><td>Sequence 2134, App</td></td>	11 <td>22.4</td> <td>74</td> <td>26</td> <td>US-10-253-471-2134</td> <td>Sequence 2134, App</td>	22.4	74	26	US-10-253-471-2134	Sequence 2134, App
33 <td>11<td>22.4</td><td>74</td><td>26</td><td>US-10-253-493-2134</td><td>Sequence 2182, App</td></td>	11 <td>22.4</td> <td>74</td> <td>26</td> <td>US-10-253-493-2134</td> <td>Sequence 2182, App</td>	22.4	74	26	US-10-253-493-2134	Sequence 2182, App
34 <td>11<td>22.4</td><td>76</td><td>1</td><td>PCT-US02-30312-2182</td><td>Sequence 2182, App</td></td>	11 <td>22.4</td> <td>76</td> <td>1</td> <td>PCT-US02-30312-2182</td> <td>Sequence 2182, App</td>	22.4	76	1	PCT-US02-30312-2182	Sequence 2182, App
35 <td>11<td>22.4</td><td>76</td><td>23</td><td>US-09-962-756-2182</td><td>Sequence 2182, App</td></td>	11 <td>22.4</td> <td>76</td> <td>23</td> <td>US-09-962-756-2182</td> <td>Sequence 2182, App</td>	22.4	76	23	US-09-962-756-2182	Sequence 2182, App
36 <td>11<td>22.4</td><td>76</td><td>26</td><td>US-10-253-471-2182</td><td>Sequence 2182, App</td></td>	11 <td>22.4</td> <td>76</td> <td>26</td> <td>US-10-253-471-2182</td> <td>Sequence 2182, App</td>	22.4	76	26	US-10-253-471-2182	Sequence 2182, App
37 <td>11<td>22.4</td><td>76</td><td>26</td><td>US-10-253-493-2182</td><td>Sequence 2182, App</td></td>	11 <td>22.4</td> <td>76</td> <td>26</td> <td>US-10-253-493-2182</td> <td>Sequence 2182, App</td>	22.4	76	26	US-10-253-493-2182	Sequence 2182, App
38 <td>11<td>22.4</td><td>13</td><td>1</td><td>PCT-US01-19843-14</td><td>Sequence 14, App1</td></td>	11 <td>22.4</td> <td>13</td> <td>1</td> <td>PCT-US01-19843-14</td> <td>Sequence 14, App1</td>	22.4	13	1	PCT-US01-19843-14	Sequence 14, App1
39 <td>10<td>20.4</td><td>13</td><td>1</td><td>PCT-US02-30312-2205</td><td>Sequence 2205, App</td></td>	10 <td>20.4</td> <td>13</td> <td>1</td> <td>PCT-US02-30312-2205</td> <td>Sequence 2205, App</td>	20.4	13	1	PCT-US02-30312-2205	Sequence 2205, App
40 <td>10<td>20.4</td><td>13</td><td>1</td><td>PCT-US02-30412-2205</td><td>Sequence 2205, App</td></td>	10 <td>20.4</td> <td>13</td> <td>1</td> <td>PCT-US02-30412-2205</td> <td>Sequence 2205, App</td>	20.4	13	1	PCT-US02-30412-2205	Sequence 2205, App
41 <td>10<td>20.4</td><td>13</td><td>1</td><td>PCT-US98-17919-6</td><td>Sequence 6, App11</td></td>	10 <td>20.4</td> <td>13</td> <td>1</td> <td>PCT-US98-17919-6</td> <td>Sequence 6, App11</td>	20.4	13	1	PCT-US98-17919-6	Sequence 6, App11
42 <td>10<td>20.4</td><td>13</td><td>12</td><td>US-08-844-462-13</td><td>Sequence 13, App1</td></td>	10 <td>20.4</td> <td>13</td> <td>12</td> <td>US-08-844-462-13</td> <td>Sequence 13, App1</td>	20.4	13	12	US-08-844-462-13	Sequence 13, App1
43 <td>10<td>20.4</td><td>13</td><td>14</td><td>US-09-011-563-9</td><td>Sequence 9, App11</td></td>	10 <td>20.4</td> <td>13</td> <td>14</td> <td>US-09-011-563-9</td> <td>Sequence 9, App11</td>	20.4	13	14	US-09-011-563-9	Sequence 9, App11
44 <td>10<td>20.4</td><td>13</td><td>19</td><td>US-09-536-556-9</td><td>Sequence 9, App11</td></td>	10 <td>20.4</td> <td>13</td> <td>19</td> <td>US-09-536-556-9</td> <td>Sequence 9, App11</td>	20.4	13	19	US-09-536-556-9	Sequence 9, App11
45 <td>10<td>20.4</td><td>13</td><td>19</td><td>US-09-536-556-9</td><td>Sequence 9, App11</td></td>	10 <td>20.4</td> <td>13</td> <td>19</td> <td>US-09-536-556-9</td> <td>Sequence 9, App11</td>	20.4	13	19	US-09-536-556-9	Sequence 9, App11

#### ALIGNMENTS

RESULT 1

US-08-973-667-4

Sequence 4, Application US/08973667

GENERAL INFORMATION:

APPLICANT: Sakakibara, Shunpei

APPLICANT: Kimura, Terutoshi

APPLICANT: Morimoto, Shigetoshi

TITLE OF INVENTION: ANTI-GLUT1-OSTEOCALCIN ANTIBODY

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: NIXON & VANDERHYTE P.C.

STREET: 1100 North Glebe Rd. 8th floor

CITY: Arlington

STATE: VA

COUNTRY: USA

ZIP: 22201-4741

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/973,667

FILING DATE: 10-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/Jp97/01246

FILING DATE: 10-APR-1997

PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: JP 8-88608
; FILING DATE: 10-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 9-43331
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 423-43
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 17
; OTHER INFORMATION: /product= "Gla"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 21
; OTHER INFORMATION: /product= "Gla"
; NAME/KEY: Modified-site
; LOCATION: 24
; OTHER INFORMATION: /product= "Gla"
; US-08-973-667-4

Query Match          93.9%; Score 46; DB 13; Length 49;
Best Local Similarity 100.0%; Pred. No. 5,2e-43;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYWMLGAPVYPDPPLKPRRYVXCLNPDCELDADHIGFQEA YRRFGYGV 49
DB 1 YLYWMLGAPVYPDPPLKPRRYVXCLNPDCELDADHIGFQEA YRRFGYGV 49

RESULT 2
US-09-036-085-5
; Sequence 5, Application US/09036085
; GENERAL INFORMATION:
; APPLICANT: Gary S. Stein et al.
; TITLE OF INVENTION: Gene Therapy Using Bone Marrow Transplants Transfected With T
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,085
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PROVISIONAL APPLICATION SERIAL NO. 60/039,839
; FILING DATE: March 6, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
```

```

; REFERENCE/DOCKET NUMBER: UMM-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-036-085-5

Query Match          51.0%; Score 25; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 9,2e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCELDADHIGFQEA YRRFGYGV 49
DB 18 LNPDCELDADHIGFQEA YRRFGYGV 42

RESULT 3
US-09-657-276-1095
; Sequence 1095, Application US/09657276
; GENERAL INFORMATION:
; APPLICANT: Conjunchem, Inc.
; APPLICANT: Bridon, Dominique
; APPLICANT: Ezrin, Alan
; APPLICANT: Milner, Peter
; APPLICANT: Holmes, Darren
; APPLICANT: Thibaudau, Karen
; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
; FILE REFERENCE: 2110
; CURRENT APPLICATION NUMBER: US/09/657,276
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; PRIOR APPLICATION NUMBER: 60/153,406
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 1617
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1095
; LENGTH: 46
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-657-276-1095

Query Match          51.0%; Score 25; DB 20; Length 46;
Best Local Similarity 100.0%; Pred. No. 1e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCELDADHIGFQEA YRRFGYGV 49
DB 22 LNPDCELDADHIGFQEA YRRFGYGV 46

RESULT 4
US-09-657-276-1093
; Sequence 1093, Application US/09657276
; GENERAL INFORMATION:
; APPLICANT: Conjunchem, Inc.
; APPLICANT: Bridon, Dominique
; APPLICANT: Ezrin, Alan
; APPLICANT: Milner, Peter
; APPLICANT: Holmes, Darren
; APPLICANT: Thibaudau, Karen
```

```

; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONUGATION TO BLOOD
; FILE REFERENCE: 2110
; CURRENT APPLICATION NUMBER: US/09/657,276
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; PRIOR APPLICATION NUMBER: 60/153,406
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 1617
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1093
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-657-276-1093

Query Match          51.0%; Score 25; DB 20; Length 47;
Best Local Similarity 100.0%; Pred. No. 1e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 23 LNPDCDELADHIGFOEAYRRFGPV 47

RESULT 5
PCT-US02-22821-108
; Sequence 108, Application PC/TUS0222821
; GENERAL INFORMATION:
; APPLICANT: HK Pharmaceuticals, Inc.
; APPLICANT: Koistler, Hubert
; APPLICANT: Siddiqui, Subaib
; APPLICANT: Little, Daniel
; TITLE OF INVENTION: Capture Compounds, Collections Thereof
; TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex
; FILE REFERENCE: 24743-2305
; CURRENT APPLICATION NUMBER: PCT/US02/22821
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 60/306,019
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/314,123
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/363,433
; PRIOR FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo Sapien
; PCT-US02-22821-108

Query Match          51.0%; Score 25; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 25 LNPDCDELADHIGFOEAYRRFGPV 49

RESULT 6
US-08-973-667-1
; Sequence 1, Application US/08973667
; GENERAL INFORMATION:
```

```

; APPLICANT: Sakakibara, Shunpei
; APPLICANT: Kimura, Terutoshi
; APPLICANT: Morimoto, Shigeto
; TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 North Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,667
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP97/01246
; FILING DATE: 10-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-88608
; FILING DATE: 10-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 9-43331
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 423-43
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 21
; OTHER INFORMATION: /product= "Gla"
; US-08-973-667-1

Query Match          51.0%; Score 25; DB 13; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 25 LNPDCDELADHIGFOEAYRRFGPV 49

RESULT 7
US-08-973-667-2
; Sequence 2, Application US/08973667
; GENERAL INFORMATION:
; APPLICANT: Sakakibara, Shunpei
; APPLICANT: Kimura, Terutoshi
; APPLICANT: Morimoto, Shigeto
; TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 North Glebe Rd. 8th floor
```

```

? CITY: Arlington
? STATE: VA
? COUNTRY: USA
? ZIP: 22201-4741
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.30
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/973,667
? FILING DATE: 10-DEC-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/JP97/01246
? FILING DATE: 10-APR-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 8-88608
? FILING DATE: 10-APR-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 9-43331
? FILING DATE: 27-FEB-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Crawford, Arthur R.
? REGISTRATION NUMBER: 25,327
? REFERENCE/DOCKET NUMBER: 423-43
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-816-4100
? TELEFAX: 703-816-4100
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 49 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? ORIGINAL SOURCE:
? ORGANISM: Human
?
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 17
? OTHER INFORMATION: /product= "Gla"
?
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 21
? OTHER INFORMATION: /product= "Gla"
?
? US-08-973-667-2
?
? Query Match 51.0%; Score 25; DB 13; Length 49;
? Best Local Similarity 100.0%; Pred. No. 1.1e-19;
? Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
?
QY 25 LNPDCELADHIGFOEAYRFGPV 49
Db 25 LNPDCELADHIGFOEAYRFGPV 49

```

```

? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.30
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/973,667
? FILING DATE: 10-DEC-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/JP97/01246
? FILING DATE: 10-APR-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 8-88608
? FILING DATE: 10-APR-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 9-43331
? FILING DATE: 27-FEB-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Crawford, Arthur R.
? REGISTRATION NUMBER: 25,327
? REFERENCE/DOCKET NUMBER: 423-43
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-816-4100
? TELEFAX: 703-816-4100
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 49 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? ORIGINAL SOURCE:
? ORGANISM: Human
?
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 17
? OTHER INFORMATION: /product= "Gla"
?
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 21
? OTHER INFORMATION: /product= "Gla"
?
? US-08-973-667-3
?
? Query Match 51.0%; Score 25; DB 13; Length 49;
? Best Local Similarity 100.0%; Pred. No. 1.1e-19;
? Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
?
QY 25 LNPDCELADHIGFOEAYRFGPV 49
Db 25 LNPDCELADHIGFOEAYRFGPV 49

```

RESULT 9

US-09-462-931-2

Sequence 2, Application US/09462931

GENERAL INFORMATION:

APPLICANT: HELLMAN, Jukka

APPLICANT: KITTEN, Sanna-Maria

APPLICANT: KARP, Matti

APPLICANT: LTVGREN, Timo

APPLICANT: VNNEN, Kalevo

APPLICANT: PETTERSSON, Kim

TITLE OF INVENTION: Isolated osteocalcin fragments

FILE REFERENCE: Isolated osteocalcin fragments

CURRENT APPLICATION NUMBER: US/09/462,931

CURRENT FILING DATE: 2000-01-18

PRIOR APPLICATION NUMBER: PCT/FI98/00550

PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: FI 973371

PRIOR FILING DATE: 1997-08-15

NUMBER OF SEQ ID NOS: 4

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 2  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: peptide  
LOCATION: (1)..(49)  
OTHER INFORMATION: Glu at residues 17, 21 and 24 may be gamma-carboxy-Glu  
US-09-462-931-2

Query Match 51.0%; Score 25; DB 18; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDDELADHIGFOEAYRRFGPV 49  
DB 25 LNPDDELADHIGFOEAYRRFGPV 49

RESULT 10  
US-10-197-954-108  
Sequence 108, Application US/10197954  
GENERAL INFORMATION:  
APPLICANT: K'ster, Hubert  
APPLICANT: Siddiqi, Suhail  
APPLICANT: Little, Daniel  
TITLE OF INVENTION: Capture Compounds, Collections Thereof  
TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex  
TITLE OF INVENTION: Compositions  
FILE REFERENCE: 24743-2305  
CURRENT APPLICATION NUMBER: US/10/197,954  
CURRENT FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: 60/306,019  
PRIOR FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/314,123  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: 60/363,433  
PRIOR FILING DATE: 2002-03-11  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 108  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-197-954-108

Query Match 51.0%; Score 25; DB 25; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDDELADHIGFOEAYRRFGPV 49  
DB 25 LNPDDELADHIGFOEAYRRFGPV 49

RESULT 11  
US-07-717-811A-5  
Sequence 5, Application US/07717811A  
GENERAL INFORMATION:  
APPLICANT: Hiroshi EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS

SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/717,811A  
FILING DATE: 19910619  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 98 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-717-811A-5

Query Match 51.0%; Score 25; DB 3; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2e-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDDELADHIGFOEAYRRFGPV 49  
DB 74 LNPDDELADHIGFOEAYRRFGPV 98

RESULT 12  
US-08-246-626-5

Sequence 5, Application US/08246626  
GENERAL INFORMATION:  
APPLICANT: Hiroshi EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 KB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,626  
FILING DATE: 20-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/717,811  
FILING DATE: 19-Jun-1991  
APPLICATION NUMBER: US 08/131,932  
FILING DATE: 05-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 98 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:

DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-246-626-5

Query Match 51.0%; Score 25; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 26-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFQEAAYRRYGPV 49  
Db 74 LNPDCDELADHIGFQEAAYRRYGPV 98

RESULT 13  
PCT-US01-08655-186  
Sequence 186, Application PC/TUS0108655  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 21272-065  
CURRENT APPLICATION NUMBER: PCT/US01/08655  
CURRENT FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: 09/522,929  
PRIOR FILING DATE: 2000-04-18  
PRIOR APPLICATION NUMBER: 09/770,160  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 09/668,317  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 09/695,783  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 09/728,628  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: 09/783,066  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 09/816,828  
PRIOR FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 584  
SOFTWARE: Custom  
SEQ ID NO 186  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-08655-186

Query Match 51.0%; Score 25; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2,1e-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFQEAAYRRYGPV 49  
Db 76 LNPDCDELADHIGFQEAAYRRYGPV 100

RESULT 14  
PCT-US01-12010-3  
Sequence 3, Application PC/TUS0112010  
GENERAL INFORMATION:  
APPLICANT: Genaisance Pharmaceuticals, Inc.  
APPLICANT: Bentivegna, Steven C.  
APPLICANT: Chew, Anne  
APPLICANT: Choi, Julie Y.  
APPLICANT: Koshy, Beena  
APPLICANT: Rounds, Eileen  
APPLICANT: Stephens, J. Claiborne  
TITLE OF INVENTION: Haplotypes of the BGLAP Gene  
FILE REFERENCE: MW-0512PCT BGLAP  
CURRENT APPLICATION NUMBER: PCT/US01/12010  
CURRENT FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 60/195,840  
PRIOR FILING DATE: 2000-04-11



; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Homo sapien  
PCT-US01-12010-3

Query Match 51.0%; Score 25; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2,1e-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCELADHIGFOEAYRRFGPY 49  
|||||  
Db 76 LNPDCELADHIGFOEAYRRFGPY 100

RESULT 15  
US-10-143-899-23  
; Sequence 23, Application US/10143899  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PTZ4/CIN  
; CURRENT APPLICATION NUMBER: US/10/143,899  
; CURRENT FILING DATE: 2002-05-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (18)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (64)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-143-899-23

Query Match 51.0%; Score 25; DB 25; Length 127;  
Best Local Similarity 100.0%; Pred. No. 2.6e-19;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCELADHIGFOEAYRRFGPY 49  
|||||  
Db 103 LNPDCELADHIGFOEAYRRFGPY 127

RESULT 16  
US-10-217-651-336  
; Sequence 336, Application US/10217651  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PZ11/CIN  
; CURRENT APPLICATION NUMBER: US/10/217,651  
; CURRENT FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: 09/760,491  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,638  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367

```
? PRIOR FILING DATE: 2000-09-29
? PRIOR APPLICATION NUMBER: 60/237,039
? PRIOR FILING DATE: 2000-10-02
? PRIOR APPLICATION NUMBER: 60/237,038
? PRIOR FILING DATE: 2000-10-02
? PRIOR APPLICATION NUMBER: 60/236,370
? PRIOR FILING DATE: 2000-09-29
? PRIOR APPLICATION NUMBER: 60/236,802
? PRIOR FILING DATE: 2000-10-02
? PRIOR APPLICATION NUMBER: 60/237,037
? PRIOR FILING DATE: 2000-10-02
? PRIOR APPLICATION NUMBER: 60/237,040
? PRIOR FILING DATE: 2000-10-02
? PRIOR APPLICATION NUMBER: 60/240,960
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/239,935
? PRIOR FILING DATE: 2000-10-13
? PRIOR APPLICATION NUMBER: 60/239,937
? PRIOR FILING DATE: 2000-10-13
? PRIOR APPLICATION NUMBER: 60/241,787
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/246,474
? PRIOR FILING DATE: 2000-11-08
? PRIOR APPLICATION NUMBER: 60/246,532
? PRIOR FILING DATE: 2000-11-08
? PRIOR APPLICATION NUMBER: 60/249,216
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,210
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/226,681
? PRIOR FILING DATE: 2000-08-22
? PRIOR APPLICATION NUMBER: 60/225,759
? PRIOR FILING DATE: 2000-08-14
? PRIOR APPLICATION NUMBER: 60/225,213
? PRIOR FILING DATE: 2000-08-14
? PRIOR APPLICATION NUMBER: 60/227,182
? PRIOR FILING DATE: 2000-08-22
? PRIOR APPLICATION NUMBER: 60/225,214
? PRIOR FILING DATE: 2000-08-14
? PRIOR APPLICATION NUMBER: 60/235,836
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: 60/230,438
? PRIOR FILING DATE: 2000-09-06
? PRIOR APPLICATION NUMBER: 60/215,135
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: 60/225,266
? PRIOR FILING DATE: 2000-08-14
? PRIOR APPLICATION NUMBER: 60/249,218
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,208
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,213
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,212
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,207
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,245
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,244
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,217
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,211
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,215
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,264
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,214
? PRIOR FILING DATE: 2000-11-17
? PRIOR APPLICATION NUMBER: 60/249,297
? PRIOR FILING DATE: 2000-11-17
```

```
? PRIOR APPLICATION NUMBER: 60/232,400
? PRIOR FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: 60/231,242
? PRIOR FILING DATE: 2000-09-08
? PRIOR APPLICATION NUMBER: 60/232,081
? PRIOR FILING DATE: 2000-09-08
? PRIOR APPLICATION NUMBER: 60/232,080
? PRIOR FILING DATE: 2000-09-08
? PRIOR APPLICATION NUMBER: 60/231,414
? PRIOR FILING DATE: 2000-09-08
? PRIOR APPLICATION NUMBER: 60/231,244
? PRIOR FILING DATE: 2000-09-08
? PRIOR APPLICATION NUMBER: 60/233,064
? PRIOR FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: 60/233,063
? PRIOR FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: 60/232,397
? PRIOR FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: 60/232,399
? PRIOR FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: 60/232,401
? PRIOR FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: 60/241,808
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/241,826
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/241,786
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/241,221
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/246,475
? PRIOR FILING DATE: 2000-11-08
? PRIOR APPLICATION NUMBER: 60/231,243
? PRIOR FILING DATE: 2000-09-08
? PRIOR APPLICATION NUMBER: 60/233,065
```

```
Query Match 51.0%; Score 25; DB 26; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.6e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 25 LNPDCDELADHIGFOEXYRRFYGV 49
DB 103 LNPDCDELADHIGFOEXYRRFYGV 127
```

```
RESULT 17
US-60-160-203-3372
; Sequence 3372, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3372
; LENGTH: 42
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-3372
```

```
Query Match 46.9%; Score 23; DB 27; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.5e-17;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 25 LNPDCDELADHIGFOEXYRRFYGV 47
DB 20 LNPDCDELADHIGFOEXYRRFYGV 42
```

```

RESULT 18
US-60-163-123-1219
: Sequence 1219, Application US/60163123
: GENERAL INFORMATION:
: APPLICANT: Bonazzoli, Vivien
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: CL000137
: CURRENT APPLICATION NUMBER: US/60/163,123
: CURRENT FILING DATE: 1999-11-02
: NUMBER OF SEQ ID NOS: 1986
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1219
: LENGTH: 42
: TYPE: PRT
: ORGANISM: Human
US-60-163-123-1219

```

Query Match	46.98;	Score 23;	DB 27;	Length 42;
Best Local Similarity	100.08;	Pred. NO. 1.5e-17;		
Matches 23; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

```
QY      25 LNPDCDELADHIGFQEA YRRFYG 47
          |||||
Db      20 LNPDCDELADHIGFQEA YRRFYG 42
```

RESULT 19  
US-60-169-840-5089  
; Sequence 5089, Application US/60169840

```

:
:
: APPLICANT: Bonazzi, Vivien
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS.
:
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
:

```

```

; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
NUMBER OF SEQ. TO NOS. 0000

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5089

```

ORGANISM: Human  
US-60-169-840-5089

Query Match	46.98;	Score 23;	DB 27;	Length 42;
Best Local Similarity	100.0%;	Pred. NO. 1.5e-17;		
Matches 23; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

```
QY      25 LNPDCDELADHIGFQEQAYRRFYG 47
         |||||
Db      20 LNPDCDELADHIGFQEQAYRRFYG 42
```

RESULT 20  
PCT-US01-08655-478

```
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
```

; CURRENT FILING DATE: 2001-04-16  
 ; PRIOR APPLICATION NUMBER: 09/522,929

PRIOR FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 09/668,317  
 ; PRIOR FILING DATE: 2000-09-22  
 ; PRIOR APPLICATION NUMBER: 09/695,783

PRIOR FILING DATE: 2000-10-24

```

1 PRIOR APPLICATION NUMBER: 09/728,628
2 PRIOR FILING DATE: 2000-12-01
3 PRIOR APPLICATION NUMBER: 09/783,066
4 PRIOR FILING DATE: 2001-02-13
5 PRIOR APPLICATION NUMBER: 09/816,828
6 PRIOR FILING DATE: 2001-03-22
7 NUMBER OF SEQ ID NOS: 584
8 SOFTWARE: Custom
9 SEQ ID NO 478
10 LENGTH: 140
11 TYPE: PRT
12 ORGANISM: Homo sapiens
13 FEATURE:
14 NAME/KEY: misc_feature
15 LOCATION: (1)...(140)
16 OTHER INFORMATION: xaa = any amino acid or nothing
17 CCT-0501-08655-478

```

Query Match	38.8%	Score 19;	DB 1;	Length 140;
Best Local Similarity	100.0%	Pred. No. 1.3e-12;		
Matches 19; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY      31 ELADHIGFQAYRRFGPV 49
        |||||
DB     122 ELADHIGFQAYRRFGPV 140

```

RESULT 21  
US-07-717-811A-10  
; sequence 10, Application US/07717811A

; APPLICANT: HIROSHI EGUCHI et al.  
 ; TITLE OF INVENTION: Recombinant Human Osteocalcin  
 ; NUMBER OF SEQUENCES: 14  
 ;

ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington

COUNTRY: U.S.A.  
ZIP: 20005

```

;
;      ADDRESS          DIRECTORY          NAME          SIZE
;-----+-----+-----+-----+-----+-----+-----+
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717, 811A
; FILING DATE: 19910619

```

```

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
;

```

NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:

TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: AMINO ACID
;

```

```

; STRANDEDNESS
; TOPOLOGY: 1
; MOLECULE TYPE:
;

```

```

;      HYPOTHETICAL:
;      ANTI-SENSE:
;      FRAGMENT TYPE:

```

ORIGINAL SOURCE:  
ORGANISM:

STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-717-811A-10

Query Match 32.7%; Score 16; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4.7e-10;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 YLYQWLGAPVPPDPL 16  
Db 1 YLYQWLGAPVPPDPL 16

RESULT 22  
US-08-246-626-10  
Sequence 10, Application US/08246626  
GENERAL INFORMATION:  
APPLICANT: Hiroshi EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,626  
FILING DATE: 20-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/717,811  
FILING DATE: 19-Jun-1991  
APPLICATION NUMBER: US 08/131,932  
FILING DATE: 05-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.

REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-246-626-10  
Query Match 32.7%; Score 16; DB 6; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4.7e-10;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 YLYQWLGAPVPPDPL 16  
Db 1 YLYQWLGAPVPPDPL 16  
RESULT 23  
US-07-717-811A-12  
Sequence 12, Application US/07717811A  
GENERAL INFORMATION:  
APPLICANT: Hiroshi EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.

COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/717,811A  
FILING DATE: 19910619  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE//DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-717-811A-12

Query Match 28.6%; Score 14; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IGFOEAYRRFGPV 49

Db 2 IGFOEAYRRFGPV 15  
|||||  
RESULT 24  
US-08-246-626-12  
; Sequence 12, Application US/08246626  
; GENERAL INFORMATION:  
; APPLICANT: Hiroshi EGUCHI et al.  
; TITLE OF INVENTION: Recombinant Human Osteocalcin  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,626  
FILING DATE: 20-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/717,811  
FILING DATE: 19-Jun-1991  
APPLICATION NUMBER: US 08/131,932  
FILING DATE: 05-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE//DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:

AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:

Query Match 28.6%; Score 14; DB 6; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 GFOEAYRRFYGPV 49  
DB 2 IGFOEAYRRFYGPV 15

RESULT 25  
US-09-341-590-44  
Sequence 44, Application US/09341590  
GENERAL INFORMATION:  
APPLICANT: Larsen, Bjarne Due  
TITLE OF INVENTION: PHARMACOLOGICALLY ACTIVE PEPTIDE CONJUGATES HAVING A  
TITLE OF INVENTION: REDUCED TENDENCY TOWARDS ENZYMATIC HYDROLYSIS  
FILE REFERENCE: PPT-20479-US  
CURRENT APPLICATION NUMBER: US/09/341,590  
CURRENT FILING DATE: 1999-07-03  
PRIOR APPLICATION NUMBER: DK 0317/98  
PRIOR FILING DATE: 1998-03-09  
NUMBER OF SEQ ID NOS: 121  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 44  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: osteocalcin (37-49)  
US-09-341-590-44

Query Match 26.5%; Score 13; DB 17; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6.5e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GFOEAYRRFYGPV 49  
DB 1 GFOEAYRRFYGPV 13

RESULT 26  
US-09-657-276-1096  
Sequence 1096, Application US/09657276  
GENERAL INFORMATION:  
APPLICANT: Conjugchem, Inc.  
APPLICANT: Bridon, Dominique  
APPLICANT: Ezrin, Alan  
APPLICANT: Milner, Peter  
APPLICANT: Holmes, Darren  
APPLICANT: Thibaudau, Karen  
TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM  
TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD  
TITLE OF INVENTION: COMPONENTS  
FILE REFERENCE: 2110  
CURRENT APPLICATION NUMBER: US/09/657,276  
CURRENT FILING DATE: 2000-09-07  
PRIOR APPLICATION NUMBER: 60/134,406  
PRIOR FILING DATE: 1999-05-17  
PRIOR APPLICATION NUMBER: 60/153,406  
PRIOR FILING DATE: 1999-09-10

PRIOR APPLICATION NUMBER: 60/159,783  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 1617  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1096  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptide  
US-09-657-276-1096

Query Match 26.5%; Score 13; DB 20; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6.5e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GFOEAYRRFYGPV 49  
DB 1 GFOEAYRRFYGPV 13

RESULT 27  
US-08-973-667-6  
Sequence 6, Application US/08973667  
GENERAL INFORMATION:  
APPLICANT: Sakakibara, Shunpei  
APPLICANT: Kimura, Terutoshi  
APPLICANT: Morimoto, Shigeto  
TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 North Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,667  
FILING DATE: 10-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP97/01246  
FILING DATE: 10-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-88608  
FILING DATE: 10-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 9-43331  
FILING DATE: 27-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 423-43  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Human  
FEATURE:

```

: NAME/KEY: Modified-site
: LOCATION: 8
: OTHER INFORMATION: /product= "Gla"
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 12
: OTHER INFORMATION: /product= "Gla"
US-08-973-667-6

Query Match      22.4%; Score 11; DB 13; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 VPPDPLKPRRXV 22
        |||
Db       1 VPPDPLKPRRXV 13

RESULT 28
US-09-036-085-22
: Sequence 22, Application US/09036085
: GENERAL INFORMATION:
: APPLICANT: Gary S. Stein et al.
: TITLE OF INVENTION: Gene Therapy Using Bone Marrow Transplants Transfected With T
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 28 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/036,085
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PROVISIONAL APPLICATION SERIAL NO. 60/039,839
: FILING DATE: March 6, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Mandragouras, Amy E.
: REGISTRATION NUMBER: 36,207
: REFERENCE/DOCKET NUMBER: UMM-007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)742-4214
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 43 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-036-085-22

Query Match      22.4%; Score 11; DB 14; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      29 CDELADHIGFQ 39
        |||
Db       22 CDELADHIGFQ 32

RESULT 29
PCT-US02-30312-2134
: Sequence 2134, Application PC/TUS0230312
: GENERAL INFORMATION:
: APPLICANT: NOVO NORDISK A/S et al.
```

```

: TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
: FILE REFERENCE: 1878-4057PC
: CURRENT APPLICATION NUMBER: PCT/US02/30312
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: 09/962,756
: PRIOR FILING DATE: 2001-09-24
: NUMBER OF SEQ ID NOS: 2227
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 2134
: LENGTH: 74
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
PCT-US02-30312-2134

Query Match      22.4%; Score 11; DB 1; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 LGAPVPPDPL 16
        |||
Db       61 LGAPVPPDPL 71

RESULT 30
PCT-US02-30412-2134
: Sequence 2134, Application PC/TUS0230412
: GENERAL INFORMATION:
: APPLICANT: NOVO NORDISK A/S et al.
: TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
: FILE REFERENCE: 1878-4056PC
: CURRENT APPLICATION NUMBER: PCT/US02/30412
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: 09/962,756
: PRIOR FILING DATE: 2001-09-24
: NUMBER OF SEQ ID NOS: 2227
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 2134
: LENGTH: 74
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
PCT-US02-30412-2134

Query Match      22.4%; Score 11; DB 1; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 LGAPVPPDPL 16
        |||
Db       61 LGAPVPPDPL 71

RESULT 31
US-09-962-756-2134
: Sequence 2134, Application US/09962756
: GENERAL INFORMATION:
: APPLICANT: PILLOTIA, RENUKA
: APPLICANT: BRISSETTE, RENE
: APPLICANT: BLOME, ARTHUR J.
: APPLICANT: SCHAEFER, LAUGE
: APPLICANT: BRANDT, JAKOB
: APPLICANT: GOLDSSTEIN, NEIL I.
: APPLICANT: SPETZLER, JANE
: APPLICANT: OSTERGAARD, SOREN
: APPLICANT: HANSEN, PER HERTZ
: TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
: FILE REFERENCE: 1878-4051US1
: CURRENT APPLICATION NUMBER: US/09/962,756
```

```
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2134
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-962-756-2134
```

```
Query Match      22.4%; Score 11; DB 23; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 LGAPVPPDPL 16
        |||||||
Db      61 LGAPVPPDPL 71
```

```
RESULT 32
US-10-253-471-2134
; Sequence 2134, Application US/10253471
; GENERAL INFORMATION:
; APPLICANT: PILDUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2134
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-471-2134
```

```
Query Match      22.4%; Score 11; DB 26; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 LGAPVPPDPL 16
        |||||||
Db      61 LGAPVPPDPL 71
```

```
RESULT 33
US-10-253-493-2134
; Sequence 2134, Application US/10253493
; GENERAL INFORMATION:
; APPLICANT: PILDUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
```

```
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2134
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-493-2134
```

```
Query Match      22.4%; Score 11; DB 26; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 LGAPVPPDPL 16
        |||||||
Db      61 LGAPVPPDPL 71
```

```
RESULT 34
PCT-US02-30312-2182
; Sequence 2182, Application PC/TUS0230312
; GENERAL INFORMATION:
; APPLICANT: NOVO NORDISK A/S et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057PC
; CURRENT APPLICATION NUMBER: PCT/US02/30312
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2182
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
PCT-US02-30312-2182
```

```
Query Match      22.4%; Score 11; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 LGAPVPPDPL 16
        |||||||
Db      63 LGAPVPPDPL 73
```

```
RESULT 35
PCT-US02-30412-2182
; Sequence 2182, Application PC/TUS0230412
; GENERAL INFORMATION:
; APPLICANT: NOVO NORDISK A/S et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056PC
; CURRENT APPLICATION NUMBER: PCT/US02/30412
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2182
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
```



PCT-US02-30412-2182

Query Match 22.4%; Score 11; DB 1; Length 76;  
Best Local Similarity 100.0%; Pred. No. 0.00056;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVPPDPL 16  
|||||  
DB 63 LGAPVPPDPL 73

RESULT 36  
US-09-962-756-2182  
; Sequence 2182, Application US/09962756

; GENERAL INFORMATION:  
; APPLICANT: PILLOTIA, RENUKA  
; APPLICANT: BRISSETTE, RENE  
; APPLICANT: BLUME, ARTHUR J.  
; APPLICANT: SCHAEFER, LAUGE  
; APPLICANT: BRANDT, JAKOB  
; APPLICANT: GOLDSTEIN, NEIL I.  
; APPLICANT: SPETZLER, JANE  
; APPLICANT: OSTERGAARD, SOREN  
; APPLICANT: HANSEN, PER HERTZ  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4051051  
; CURRENT APPLICATION NUMBER: US/09/962,756  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/538,038  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 09/146,127  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 2182  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-962-756-2182

Query Match 22.4%; Score 11; DB 23; Length 76;  
Best Local Similarity 100.0%; Pred. No. 0.00056;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVPPDPL 16  
|||||  
DB 63 LGAPVPPDPL 73

RESULT 37  
US-10-253-471-2182  
; Sequence 2182, Application US/10253471  
; GENERAL INFORMATION:  
; APPLICANT: PILLOTIA, RENUKA et al.  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4057  
; CURRENT APPLICATION NUMBER: US/10/253,471  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: 09/962,756  
; PRIOR FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/538,038  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 09/146,127  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 2182  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-10-253-471-2182

Query Match 22.4%; Score 11; DB 26; Length 76;  
Best Local Similarity 100.0%; Pred. No. 0.00056;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVPPDPL 16  
|||||  
DB 63 LGAPVPPDPL 73

RESULT 38  
US-10-253-493-2182  
; Sequence 2182, Application US/10253493  
; GENERAL INFORMATION:  
; APPLICANT: PILLOTIA, RENUKA et al.  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4056  
; CURRENT APPLICATION NUMBER: US/10/253,493  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: 09/962,756  
; PRIOR FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/538,038  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 09/146,127  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 2182  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-10-253-493-2182

Query Match 22.4%; Score 11; DB 26; Length 76;  
Best Local Similarity 100.0%; Pred. No. 0.00056;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVPPDPL 16  
|||||  
DB 63 LGAPVPPDPL 73

RESULT 39  
PCT-US00-19843-14  
; Sequence 14, Application PC/TUS0019843  
; GENERAL INFORMATION:  
; APPLICANT: Herr, John C.  
; APPLICANT: Norton, Elizabeth J.  
; APPLICANT: Deikman, Alan B.  
; TITLE OF INVENTION: Recombinant Antibody Directed Against Human Sperm  
; TITLE OF INVENTION: Antigen  
; FILE REFERENCE: 00415-02  
; CURRENT APPLICATION NUMBER: PCT/US00/19843  
; CURRENT FILING DATE: 2000-07-21  
; PRIOR APPLICATION NUMBER: 60/145,512  
; PRIOR FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: commercially  
; OTHER INFORMATION: available peptide antigen  
PCT-US00-19843-14

Query Match 20.4%; Score 10; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDPL 16  
|||||  
DB 1 GAVPYPPDPL 10

## RESULT 40

PCT-US02-30312-2205  
; Sequence 2205, Application PC/TUS0230312  
; GENERAL INFORMATION:  
; APPLICANT: NOVO NORDISK A/S et al.  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4057PC  
; CURRENT APPLICATION NUMBER: PCT/US02/30312  
; PRIOR FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: 09/962,756  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2205  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
PCT-US02-30312-2205

Query Match 20.4%; Score 10; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDPL 16  
|||||  
DB 1 GAVPYPPDPL 10

## RESULT 41

PCT-US02-30412-2205  
; Sequence 2205, Application PC/TUS0230412  
; GENERAL INFORMATION:  
; APPLICANT: NOVO NORDISK A/S et al.  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4056PC  
; CURRENT APPLICATION NUMBER: PCT/US02/30412  
; PRIOR FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: 09/962,756  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2205  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
PCT-US02-30412-2205

Query Match 20.4%; Score 10; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDPL 16  
|||||  
DB 1 GAVPYPPDPL 10

## RESULT 42

PCT-US98-17919-6  
; Sequence 6, Application PC/TUS9817919  
; GENERAL INFORMATION:  
; APPLICANT: Amgen Inc.  
; TITLE OF INVENTION: A FIBROBLAST GROWTH FACTOR  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Amgen Inc.  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91320-1789

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US98/17919  
FILING DATE:  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Mazza, Richard J.  
REGISTRATION NUMBER: 27,657  
REFERENCE/DOCKET NUMBER: A-469  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805.447.4112  
TELEFAX: 805.447.1090  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US98-17919-6

Query Match 20.4%; Score 10; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDPL 16  
|||||  
DB 1 GAVPYPPDPL 10

## RESULT 43

US-08-844-462-13  
; Sequence 13, Application US/08844462  
; GENERAL INFORMATION:  
; APPLICANT: PATTERSON, DALE H.  
; TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Administrator - Testa, Hurwitz &  
; STREET: High Street Tower, 125 High Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/844,462  
FILING DATE:  
CLASSIFICATION:

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/796,598
; FILING DATE: 07-FEB-1997
; APPLICATION NUMBER: US 08/446,055
; FILING DATE: 19-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FLYNN Esq., Kerry A.
; REGISTRATION NUMBER: 33,693
; REFERENCE/DOCKET NUMBER: STP-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-844-462-13

Query Match                20.4%; Score 10; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 GAPVYPDPPL 16
    |||
Db 1 GAPVYPDPPL 10
```

```

RESULT 44
US-09-011-563-9
; Sequence 9, Application US/09011563
; GENERAL INFORMATION:
; APPLICANT: Wang, Linfa
; TITLE OF INVENTION: EPI TOPE TAGGING SYSTEM
; FILE REFERENCE: Griffith Hack
; CURRENT APPLICATION NUMBER: US/09/011,563
; CURRENT FILING DATE: 1998-06-15
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide
; OTHER INFORMATION: sequence
US-09-011-563-9
```

```

Query Match                20.4%; Score 10; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 GAPVYPDPPL 16
    |||
Db 1 GAPVYPDPPL 10
```

```

RESULT 45
US-09-536-556-9
; Sequence 9, Application US/09536556
; GENERAL INFORMATION:
; APPLICANT: Wang, Linfa
; TITLE OF INVENTION: EPI TOPE TAGGING SYSTEM
; FILE REFERENCE: Griffith Hack
; CURRENT APPLICATION NUMBER: US/09/536,556
; CURRENT FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 09/011,563
; PRIOR FILING DATE: 1998-06-15
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
```

```

; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide
; OTHER INFORMATION: sequence
US-09-536-556-9
```

```

Query Match                20.4%; Score 10; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 GAPVYPDPPL 16
    |||
Db 1 GAPVYPDPPL 10
```

```

Search completed: December 4, 2002, 15:53:49
Job time : 141 secs
```





```
; Patent No. 5164483
; APPLICANT: Takashi, Kurihara, Elji, Taniyama, Sachio, Hirose
; TITLE OF INVENTION: Y-CARBOXYGLUTAMATE DERIVATIVE METHOD
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN
; USING THE SAME
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/575,639
; FILING DATE: 31-AUG-1990
; SEQ ID NO:2:
; LENGTH: 49
5164483-2

Query Match          61.2%; Score 30; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 2e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 PRRVCXLPDCDELADHIGFOEAYRRFGPV 49
Db 18 PRRVCXLPDCDELADHIGFOEAYRRFGPV 49

RESULT 4
5434245-1
; Patent No. 5434245
; APPLICANT: KOYAMA, NOBUTO, KIMIZUKA, FUSAO, KATO, IKUNOSHIN
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING
; THE SAME
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/993,980
; FILING DATE: 16-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 855,473; 444,786
; FILING DATE: 23-MAR-1992; 01-DEC-1989
; APPLICATION NUMBER: 444,786
; FILING DATE: 01-DEC-1989
; SEQ ID NO:1:
; LENGTH: 49
5434245-1

Query Match          51.0%; Score 25; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 25 LNPDCDELADHIGFOEAYRRFGPV 49

RESULT 5
5434245-2
; Patent No. 5434245
; APPLICANT: KOYAMA, NOBUTO, KIMIZUKA, FUSAO, KATO, IKUNOSHIN
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING
; THE SAME
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/993,980
; FILING DATE: 16-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 855,473; 444,786
; FILING DATE: 23-MAR-1992; 01-DEC-1989
; APPLICATION NUMBER: 444,786
; FILING DATE: 01-DEC-1989
; SEQ ID NO:2:
; LENGTH: 50
5434245-2

Query Match          51.0%; Score 25; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.5e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 25 LNPDCDELADHIGFOEAYRRFGPV 49

RESULT 6
5434245-3
; Patent No. 5434245
; APPLICANT: KOYAMA, NOBUTO, KIMIZUKA, FUSAO, KATO, IKUNOSHIN
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING
; THE SAME
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/993,980
; FILING DATE: 16-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 855,473; 444,786
; FILING DATE: 23-MAR-1992; 01-DEC-1989
; APPLICATION NUMBER: 444,786
; FILING DATE: 01-DEC-1989
; SEQ ID NO:3:
; LENGTH: 50
5434245-3

Query Match          51.0%; Score 25; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.5e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 26 LNPDCDELADHIGFOEAYRRFGPV 50

RESULT 7
5434245-4
; Patent No. 5434245
; APPLICANT: KOYAMA, NOBUTO, KIMIZUKA, FUSAO, KATO, IKUNOSHIN
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING
; THE SAME
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/993,980
; FILING DATE: 16-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 855,473; 444,786
; FILING DATE: 23-MAR-1992; 01-DEC-1989
; APPLICATION NUMBER: 444,786
; FILING DATE: 01-DEC-1989
; SEQ ID NO:4:
; LENGTH: 51
5434245-4

Query Match          51.0%; Score 25; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 4.6e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49
Db 26 LNPDCDELADHIGFOEAYRRFGPV 50

RESULT 8
5168041-1
; Patent No. 5168041
; APPLICANT: BERGMANN, ANDREAS E.
; TITLE OF INVENTION: METHOD FOR THE DETERMINATION OF
; OSTEOCALCIN IN HUMAN SERUM OR PLASMA
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/416,728
; FILING DATE: 03-OCT-1989
; SEQ ID NO:1:
; LENGTH: 48
```

5168041-1

Query Match 32.7%; Score 16; DB 6; Length 48;

Best Local Similarity 100.0%; Pred. No. 1.0e-10;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDL 16

DB 1 YLYQWLGAPVPPDL 16

RESULT 9

US-08-796-598-13

Sequence 13, Application US/08796598

Patent No. 5827659

GENERAL INFORMATION:

APPLICANT: PATTERSON, DALE H.

APPLICANT: TARR, GEORGE E.

TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING

TITLE OF INVENTION: POLYMERS USING MASS SPECTROMETRY.

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Administrator - Testa, Hurwitz &amp;

ADDRESSEE: Thibault

STREET: High Street Tower, 125 High Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/796,598

FILING DATE: 07-FEB-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/446,055

FILING DATE: 19-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: FLYNN Esq., Kerry A.

REGISTRATION NUMBER: 33,693

REFERENCE/DOCKET NUMBER: SYP-115

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-796-598-13

Query Match 20.4%; Score 10; DB 2; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVPPDL 16

DB 1 GAPVPPDL 10

RESULT 10

US-08-447-175A-13

Sequence 13, Application US/08447175A

Patent No. 5869240

GENERAL INFORMATION:

APPLICANT: PATTERSON, DALE H.

TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING

TITLE OF INVENTION: POLYMERS WITH A STATISTICAL CERTAINTY USING MASS

SPECTROMETRY.

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Administrator - Testa, Hurwitz &amp;

ADDRESSEE: Thibault, LLP

STREET: High Street Tower, 125 High Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/447,175A

FILING DATE: 19-MAY-1995

CLASSIFICATION: 422

ATTORNEY/AGENT INFORMATION:

NAME: RAUSCHENBACH, Kurt

REGISTRATION NUMBER: 40,137

REFERENCE/DOCKET NUMBER: SYP-114

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-447-175A-13

Query Match 20.4%; Score 10; DB 2; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVPPDL 16

DB 1 GAPVPPDL 10

RESULT 11

US-08-943-915-6

Sequence 6, Application US/08943915

Patent No. 5998170

GENERAL INFORMATION:

APPLICANT: Itoh, No. 5998170uyuki

APPLICANT: Martin, Frank

TITLE OF INVENTION: A FIBROBLAST GROWTH FACTOR

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: 1840 DeHavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/943,915

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mazza, Richard J.

REGISTRATION NUMBER: 27,657  
REFERENCE/DOCKET NUMBER: A-469  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805.447.4112  
TELEFAX: 805.447.1090  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-943-915-6

Query Match 20.4%; Score 10; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVYPDPL 16  
|||||  
Db 1 GAPVYPDPL 10

RESULT 12  
US-08-881-037-112  
Sequence 112, Application US/08881037  
Patent No. 6080588  
GENERAL INFORMATION:  
APPLICANT: Glick, Gary D.  
APPLICANT: Swanson, Patrick C.  
TITLE OF INVENTION: DNA BINDING ANTIBODIES  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/881,037  
FILING DATE: 23-JUN-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/443,540  
FILING DATE: 18-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Konski, Antoinette F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: 203442110710  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 813-5600  
TELEFAX: (650) 494-0792  
TELEX:  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-881-037-112

Query Match 20.4%; Score 10; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAPVYPDPL 16  
|||||

|||||  
Db 1 GAPVYPDPL 10

RESULT 13  
US-08-652-816A-43  
Sequence 43, Application US/08652816A  
Patent No. 5872215  
GENERAL INFORMATION:  
APPLICANT: Osbourn, JK  
APPLICANT: Allen, DJ  
APPLICANT: McCarter, JG  
TITLE OF INVENTION: Specific binding members, materials and  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,816A  
FILING DATE: 23-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125579.4  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125579.8  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206318.9  
FILING DATE: 24-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206372.6  
FILING DATE: 23-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9525004.9  
FILING DATE: 07-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610824.6  
FILING DATE: 23-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/02240  
FILING DATE: 02-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/244,597  
FILING DATE: 01-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/33308  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-652-816A-43

Query Match 20.4%; Score 10; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAPVYPDPL 16  
|||||



Db 23 GAPVPDPL 32

## RESULT 14

US-08-466-120-2

; Sequence 2, Application US/08466120

; Patent No. 5869284

## GENERAL INFORMATION:

APPLICANT: CAO, ET AL.

TITLE OF INVENTION: Retinoic Acid Receptor Epsilon

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART &amp; OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,120

FILING DATE: June 6, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA: PCT/US94/07266

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN

US-08-466-120-2

Query Match 14.3%; Score 7; DB 2; Length 433;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WLGAPVP 11

Db 4 WLGAPVP 10

## RESULT 15

PCT-US94-07266-2

; Sequence 2, Application PC/US9407266

## GENERAL INFORMATION:

APPLICANT: CAO, ET AL.

TITLE OF INVENTION: Retinoic Acid Receptor Epsilon

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART &amp; OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

## COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/07266

FILING DATE: Concurrently

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN

PCT-US94-07266-2

Query Match 14.3%; Score 7; DB 5; Length 433;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WLGAPVP 11

Db 4 WLGAPVP 10

## RESULT 16

US-08-333-358-8

; Sequence 8, Application US/08333358

Patent No. 5571696

## GENERAL INFORMATION:

APPLICANT: EVANS Ph.D., RONALD M.

APPLICANT: MANGELSDORF Ph.D., DAVID J.

APPLICANT: ONG Ms., ESTELITA S.

APPLICANT: ORO Ph.D., ANTHONY E.

APPLICANT: BORGMEYER Ph.D., UWE K.

APPLICANT: GIGUERE Ph.D., VINCENT NMN

TITLE OF INVENTION: NOVEL RECEPTORS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pretty, Schroeder, Brueggemann &amp; Clark

STREET: 444 So. Flower St., Suite 2000

CITY: Los Angeles

STATE: CA

COUNTRY: US

ZIP: 90071-2921

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/333,358

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/761,068

ATTORNEY/AGENT INFORMATION:

NAME: Reiter Ph.D., Stephen E.

REGISTRATION NUMBER: 31192

REFERENCE/DOCKET NUMBER: P31 8936

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 440 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-333-358-8

Query Match 14.3%; Score 7; DB 1; Length 440;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WIGAPVP 11  
|||||  
DB 4 WIGAPVP 10

RESULT 17  
US-08-463-694-8  
Sequence 8, Application US/08463694  
Patent No. 5686233

GENERAL INFORMATION:  
APPLICANT: EVANS Ph.D., RONALD M.  
APPLICANT: MANGELSDORF Ph.D., DAVID J.  
APPLICANT: ONG Ms., ESTELITA S.  
APPLICANT: ORO Ph.D., ANTHONY E.  
APPLICANT: BORGMEYER Ph.D., UWE K.  
APPLICANT: GIGUERE Ph.D., VINCENT NMN  
APPLICANT: YAO Mr., TSO-PANG NMN  
TITLE OF INVENTION: NOVEL RECEPTORS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
STREET: 444 So. Flower St., Suite 2000  
CITY: Los Angeles  
STATE: CA  
COUNTRY: US  
ZIP: 90071-2921

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,694  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/761,068  
FILING DATE: 17-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter Ph.D., Stephen E.  
REGISTRATION NUMBER: 31192  
REFERENCE/DOCKET NUMBER: P31 8936  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 440 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-694-8

Query Match 14.3%; Score 7; DB 1; Length 440;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WIGAPVP 11  
|||||

DB 4 WIGAPVP 10

RESULT 18  
US-08-694-501-8  
Sequence 8, Application US/08694501  
Patent No. 5710004

GENERAL INFORMATION:  
APPLICANT: EVANS Ph.D., RONALD M.  
APPLICANT: MANGELSDORF Ph.D., DAVID J.  
APPLICANT: ONG Ms., ESTELITA S.  
APPLICANT: ORO Ph.D., ANTHONY E.  
APPLICANT: BORGMEYER Ph.D., UWE K.  
APPLICANT: GIGUERE Ph.D., VINCENT NMN  
APPLICANT: YAO Mr., TSO-PANG NMN  
TITLE OF INVENTION: NOVEL RECEPTORS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
STREET: 444 So. Flower St., Suite 2000  
CITY: Los Angeles  
STATE: CA  
COUNTRY: US  
ZIP: 90071-2921

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/694,501  
FILING DATE: 07-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/333,358  
FILING DATE:  
APPLICATION NUMBER: US/07/761,068  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter Ph.D., Stephen E.  
REGISTRATION NUMBER: 31192  
REFERENCE/DOCKET NUMBER: P31 8936  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 440 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-501-8

Query Match 14.3%; Score 7; DB 1; Length 440;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WIGAPVP 11  
|||||  
DB 4 WIGAPVP 10

RESULT 19  
US-08-373-935-1  
Sequence 1, Application US/08373935  
Patent No. 5747661

GENERAL INFORMATION:  
APPLICANT: Evans, Ronald M.  
APPLICANT: Mangelsdorf, David J.  
APPLICANT: Willy, Patricia J.  
TITLE OF INVENTION: IDENTIFICATION OF A DISTINCT  
RETINOID-RESPONSIVE PATHWAY AND USES THEREFOR  
NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
; STREET: 444 South Flower Street, Suite 2000  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/373,935  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Reiter, Stephen E.  
; REGISTRATION NUMBER: 31,192  
; REFERENCE/DOCKET NUMBER: P41 9894  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-546-4737  
; TELEFAX: 619-546-4737  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-373-935-1

Query Match 14.3% Score 7; DB 1;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WLGAVPV 11  
Db 4 WLGAVPV 10

RESULT 20  
US-08-968-747-7  
; Sequence 7, Application US/08968747  
; Patent No. 6060595  
; GENERAL INFORMATION:  
; APPLICANT: Scaglioni et al.  
; TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/968,747  
; FILING DATE: 03-SEP-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 08472/705001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-968-747-7

Query Match 12.2% Score 6; DB 3;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 GAPVPY 12  
Db 1 GAPVPY 6

RESULT 21  
US-08-493-071-28  
; Sequence 28, Application US/08493071  
; Patent No. 6127149  
; GENERAL INFORMATION:  
; APPLICANT: Hirai, Yohei  
; APPLICANT: Koshida, Shogo  
; APPLICANT: Oka, Yumiko  
; TITLE OF INVENTION: MODIFIED EPIMORPHIN  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER  
; STREET: 99 CANAL CENTER PLAZA, SUITE 300  
; CITY: ALEXANDRIA  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22314  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/493,071  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Price, Robert L.  
; REGISTRATION NUMBER: 22,685  
; REFERENCE/DOCKET NUMBER: 715-107  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-684-1111  
; TELEFAX: 703-684-1124  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-493-071-28

Query Match 12.2% Score 6; DB 3;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 GAPVPY 12  
Db 1 GAPVPY 6

RESULT 22  
US-08-791-522-4  
; Sequence 4, Application US/08791522  
; Patent No. 5935817  
; GENERAL INFORMATION:

APPLICANT: Bandman, Olga  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/791,522  
FILING DATE: Filed Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0193 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 118195  
US-08-791-522-4

Query Match 12.2%; Score 6; DB 2; Length 139;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11  
|||||  
DB 31 LGAPVP 36

RESULT 23  
US-09-314-777-4  
Sequence 4, Application US/09314777  
Patent No. 6110686  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/314,777  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/791,522  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0193 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 118195  
US-09-314-777-4

Query Match 12.2%; Score 6; DB 3; Length 139;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11  
|||||  
DB 31 LGAPVP 36

RESULT 24  
US-08-457-176-2  
Sequence 2, Application US/08457176  
Patent No. 5591826  
GENERAL INFORMATION:  
APPLICANT: Vogelstein, Bert  
APPLICANT: Kinzler, Kenneth W.  
APPLICANT: de la Chapelle, Albert  
TITLE OF INVENTION: Mutator Gene and Hereditary  
TITLE OF INVENTION: No. 5591826-Polypsoid Colorectal Cancer  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie, and Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,176  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160295  
FILING DATE: 02-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.44900  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202.508.9100  
TELEFAX: 202.508.9299  
TELEX: 197430 BBMB UT  
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 934 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-457-175-2

Query Match 12.2%; Score 6; DB 1; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVPY 12  
DB 614 GAPVPY 619

RESULT 25  
US-08-457-175-2  
Sequence 2, Application US/08457175  
Patent No. 5693470  
GENERAL INFORMATION:  
APPLICANT: Vogelstein, Bert  
APPLICANT: Kinzler, Kenneth W.  
TITLE OF INVENTION: Mutator Gene and Hereditary  
TITLE OF INVENTION: No. 5693470-Polypsosis colorectal Cancer  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie, and Beckelt  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,175  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/160295  
FILING DATE: 02-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.44900  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202.508.9100  
TELEFAX: 202.508.9299  
TELEX: 197430 BBMB UT  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 934 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-457-175-2

Query Match 12.2%; Score 6; DB 1; Length 934;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVPY 12  
DB 614 GAPVPY 619

RESULT 26  
US-08-709-784-1  
Sequence 1, Application US/08709784  
Patent No. 6048701  
GENERAL INFORMATION:  
APPLICANT: The Johns Hopkins University  
TITLE OF INVENTION: Antibody Detection of Mismatch Repair  
TITLE OF INVENTION: Protein  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti, Ltd.  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20001-4597  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/709,784  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/480,351  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.57434  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 934 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-709-784-1

Query Match 12.2%; Score 6; DB 3; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVPY 12  
DB 614 GAPVPY 619

RESULT 27  
US-09-651-656-3  
Sequence 3, Application US/09651656  
Patent No. 6340566  
GENERAL INFORMATION:  
APPLICANT: MCCUTCHEN-MALONEY, SANDRA  
APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY  
TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE

```

; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES
; FILE REFERENCE: IL-10689
; CURRENT APPLICATION NUMBER: US/09/651,656
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 934
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-651-656-3

Query Match      12.2%; Score 6; DB 4; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 GAPVPY 12
        |||||
DB      614 GAPVPY 619

RESULT 28
US-09-650-855-3
; Sequence 3, Application US/09650855
; Patent No. 6365355
; GENERAL INFORMATION:
; APPLICANT: MCOTHEEN-MALONEY, SANDRA
; TITLE OF INVENTION: LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA
; FILE REFERENCE: IL-10284
; CURRENT APPLICATION NUMBER: US/09/650,855
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 934
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-650-855-3

Query Match      12.2%; Score 6; DB 4; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 GAPVPY 12
        |||||
DB      614 GAPVPY 619

RESULT 29
US-07-642-734C-5
; Sequence 5, Application US/07642734C
; Patent No. 5824513
; GENERAL INFORMATION:
; APPLICANT: Katz, L
; APPLICANT: Donadio, S
; APPLICANT: McAlpine, J B
; TITLE OF INVENTION: Recombinant DNA Method for Producing
; TITLE OF INVENTION: Erythromycin Analogs
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
```

```

; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,734C
; FILING DATE: 17-JAN-91
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckerts, Andreas M
; REGISTRATION NUMBER: 32652
; REFERENCE/DOCKET NUMBER: 4952.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-9396
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3170 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-642-734C-5

Query Match      12.2%; Score 6; DB 2; Length 3170;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      32 LADHIG 37
        |||||
DB      2884 LADHIG 2889

RESULT 30
US-08-439-009A-5
; Sequence 5, Application US/08439009A
; Patent No. 6004787
; GENERAL INFORMATION:
; APPLICANT: Donadio, S
; APPLICANT: Katz, L
; APPLICANT: McAlpine, J B
; TITLE OF INVENTION: Method of Directing Biosynthesis of
; TITLE OF INVENTION: Specific Polyketides
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven F. Weinstock
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,009A
; FILING DATE: 11-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Casulo, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 4952.US.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3170 amino acids
; TYPE: amino acid
```

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-439-009A-5

Query Match  
Best Local Similarity 12.2%; Score 6; DB 3; Length 3170;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LADHIG 37  
DB 2884 LADHIG 2889

RESULT 31  
US-08-991-789A-138  
Sequence 138, Application US/08991789A  
Patent No. 6225054  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 292  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed IP Law Group  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,789A  
FILING DATE: 11-Dec-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potler, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-991-789A-138

Query Match  
Best Local Similarity 10.2%; Score 5; DB 4; Length 9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
DB 3 QEAYR 7

RESULT 32  
US-09-062-451-138  
Sequence 138, Application US/09062451  
Patent No. 6344550  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
APPLICANT: Reed, Steven G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 297  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/062,451  
FILING DATE: 04-Apr-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.419C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-062-451-138

Query Match  
Best Local Similarity 10.2%; Score 5; DB 4; Length 9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
DB 3 QEAYR 7

RESULT 33  
US-09-598-326-138  
Sequence 138, Application US/09598326  
Patent No. 6423496  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 247  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed Intellectual Property Law Group PLLC  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/598,326  
FILING DATE: 20-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potler, Jane E. R.  
REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 210121.419D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 138:  
US-09-598-326-138

Query Match 10.2%; Score 5; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
|||||  
Db 3 QEAYR 7

RESULT 34  
US-09-142-974B-5  
Sequence 5, Application US/09142974B  
Patent No. 6451995  
GENERAL INFORMATION:  
APPLICANT: Cheung, Nai-Kong V.  
APPLICANT: Larson, Steven M.  
APPLICANT: Guo, Hong-Fen  
APPLICANT: Rivlin, Ken  
APPLICANT: Sadelaï, Michel  
TITLE OF INVENTION: Single Chain Fv Constructs of Anti-Ganglioside GD2  
FILE REFERENCE: MSK-P-013-USNP  
CURRENT APPLICATION NUMBER: US/09/142,974B  
CURRENT FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: PCT/US97/04427  
PRIOR FILING DATE: 1997-03-20  
PRIOR APPLICATION NUMBER: 60/013,703  
PRIOR FILING DATE: 1996-03-20  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: antibody tag  
US-09-142-974B-5

Query Match 10.2%; Score 5; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVP 11  
|||||  
Db 1 GAVP 5

RESULT 35  
PCT-US93-05647-18  
Sequence 18, Application PC/TUS9305647  
GENERAL INFORMATION:  
APPLICANT: President and Fellows of Harvard College and Eunice Kennedy Shriver C  
TITLE OF INVENTION: HETEROGENEOUS PROTEOLIPID PEPTIDE  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: 2 Militia Drive  
CITY: Lexington  
STATE: MA

ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/05647  
FILING DATE: 19930610  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/896,704  
FILING DATE: JUNE 10, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: H092-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-9540  
TELEFAX: (617) 861-6240  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US93-05647-18

Query Match 10.2%; Score 5; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WLGP 9  
|||||  
Db 6 WLGP 10

RESULT 36  
US-08-101-041A-4  
Sequence 4, Application US/08101041A  
Patent No. 5541101  
GENERAL INFORMATION:  
APPLICANT: Saji, Fumitaka  
APPLICANT: AZUMA, Chihito  
APPLICANT: KIMURA, Tadashi  
TITLE OF INVENTION: ANTI-OXYTOCIN RECEPTOR ANTIBODIES AND  
NUMBER OF SEQUENCES: 7  
METHODS FOR THEIR PRODUCTION  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P. O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/101,041A  
FILING DATE: 03-AUG-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-206854  
FILING DATE: 03-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 002258-008  
TELECOMMUNICATION INFORMATION:



TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..18  
OTHER INFORMATION: /note="Amino acids 102-119 of the  
OTHER INFORMATION: oxytocin receptor polypeptide."  
US-08-101-041A-4

Query Match 10.2%; Score 5; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 RYGP 48  
|||||  
DB 3 RYGP 7

RESULT 37  
US-08-602-999A-125  
Sequence 125, Application US/08602999A  
Patent No. 6184205  
GENERAL INFORMATION:  
APPLICANT: SPARKS, Andrew B.  
APPLICANT: KAY, Brian K.  
APPLICANT: THORN, Judith M.  
APPLICANT: QUILIAM, Lawrence A.  
APPLICANT: DER, Channing J.  
APPLICANT: FOWLKES, Dana M.  
APPLICANT: RIDER, James E.  
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,999A  
FILING DATE: 16-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MISTOCK, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 125:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-602-999A-125  
Query Match 10.2%; Score 5; DB 4; Length 20;

Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVP 11  
|||||  
DB 7 GAPVP 11

RESULT 38  
US-09-500-124-125  
Sequence 125, Application US/09500124  
Patent No. 6432920  
GENERAL INFORMATION:  
APPLICANT: SPARKS, Andrew B.  
APPLICANT: KAY, Brian K.  
APPLICANT: THORN, Judith M.  
APPLICANT: QUILIAM, Lawrence A.  
APPLICANT: DER, Channing J.  
APPLICANT: FOWLKES, Dana M.  
APPLICANT: RIDER, James E.  
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/500,124  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/602,999  
FILING DATE: 16-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MISTOCK, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 125:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-500-124-125  
Query Match 10.2%; Score 5; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVP 11  
|||||  
DB 7 GAPVP 11

RESULT 39  
US-08-403-378B-12  
Sequence 12, Application US/08403378B  
Patent No. 5759991  
GENERAL INFORMATION:  
APPLICANT: TOHDOH, NAOKI

APPLICANT: TOJO, SHIN-ICHIRO  
APPLICANT: KOJIMA, SHIN-ICHI  
APPLICANT: Ueki, YASUYUKI  
APPLICANT: NISHIHARA, TOSHIO  
APPLICANT: FUKUSHIMA, NOBUYUKI  
APPLICANT: IRIE, TSUNEMASA  
APPLICANT: ONO, KEIICHI  
APPLICANT: AGUI, HIDEO  
APPLICANT: OJIKI, KOSEI  
TITLE OF INVENTION: NEUROTROPHIC PEPTIDE DERIVATIVES  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 PENNSYLVANIA AVENUE, NW  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20037-3202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,378B  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-124688  
FILING DATE: 27-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-080398  
FILING DATE: 30-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-280590  
FILING DATE: 27-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-333241  
FILING DATE: 21-DEC-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-243003  
FILING DATE: 12-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/758,043  
FILING DATE: 12-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/873,764  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP93/01214  
FILING DATE: 27-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGART, WADDELL, A  
REGISTRATION NUMBER: 24,861  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
TELEX: 6491103  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: rattus norvegicus  
STRAIN: Wistar  
TISSUE TYPE: hippocampal tissue of brain  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..22

US-08-403-378B-12  
Query Match 10.2%; Score 5; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 LGAPV 10  
|||||  
Db 3 LGAPV 7  
RESULT 40  
US-08-991-789A-133  
Sequence 133; Application US/08991789A  
Patent No. 6225054  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 292  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed IP Law Group  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,789A  
FILING DATE: 11-Dec-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 133:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 133:  
US-08-991-789A-133  
Query Match 10.2%; Score 5; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 QEAYR 43  
|||||  
Db 10 QEAYR 14  
RESULT 41  
US-09-062-451-133  
Sequence 133; Application US/09062451  
Patent No. 6344550  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER

NUMBER OF SEQUENCES: 297  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/062,451  
FILING DATE: 04-APR-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: MAKI, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.419C2.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 133:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-062-451-133

Query Match 10.2%; Score 5; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 QEAYR 43  
|||||  
Db 10 QEAYR 14

RESULT 42  
US-09-598-326-133  
Sequence 133, Application US/09598326  
Patent No. 6423496  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 247  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed Intellectual Property Law Group PLLC  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/598,326  
FILING DATE: 20-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potler, Jane E.R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419D1  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 133:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 133:  
US-09-598-326-133

Query Match 10.2%; Score 5; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 QEAYR 43  
|||||  
Db 10 QEAYR 14

RESULT 43  
US-08-324-977-42  
Sequence 42, Application US/08324977  
Patent No. 5747339  
GENERAL INFORMATION:  
APPLICANT: OKAYAMA, Hiroto  
APPLICANT: FUKU, Isao  
APPLICANT: MORI, Chisato  
APPLICANT: TAKAMIZAWA, Akahisa  
APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westernman, Hattori, Mclelland &  
ADDRESS: Naughton  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/324,977  
FILING DATE: 18-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-167466  
FILING DATE: 25-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-230921  
FILING DATE: 31-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-305605  
FILING DATE: 09-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/099,706  
FILING DATE: 30-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/769,996  
FILING DATE: 02-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/635,451  
FILING DATE: 28-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens-Smith, Theresa M.  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: 900703D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930

TELEFAX: (202) 887-0357  
TELEX: 440142  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-324-977-42

Query Match 10.2%; Score 5; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELA 33  
|||||  
DB 16 CDELA 20

RESULT 44  
US-08-384-616-42  
Sequence 42, Application US/08384616  
Patent No. 5647101  
GENERAL INFORMATION:  
APPLICANT: OKAYAMA, Hiroto  
APPLICANT: FUKE, Isao  
APPLICANT: MORI, Chisato  
APPLICANT: TAKAMIZAWA, Akahisa  
APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/384,616  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/769,996  
FILING DATE: 02-OCT-1991  
APPLICATION NUMBER: JP 2-167466  
FILING DATE: 25-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-230921  
FILING DATE: 31-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-305605  
FILING DATE: 09-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/635,451  
FILING DATE: 28-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens-Smith, Theresa M.  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: 900703B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930  
TELEFAX: (202) 887-0357  
TELEX: 440142  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-384-616-42

Query Match 10.2%; Score 5; DB 2; Length 30;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELA 33  
|||||  
DB 16 CDELA 20

RESULT 45  
US-08-904-686A-42  
Sequence 42, Application US/08904686A  
Patent No. 5998130  
GENERAL INFORMATION:  
APPLICANT: OKAYAMA, Hiroto  
APPLICANT: FUKE, Isao  
APPLICANT: MORI, Chisato  
APPLICANT: TAKAMIZAWA, Akahisa  
APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/904,686A  
FILING DATE: 01-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/324,977  
FILING DATE: 18-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-167466  
FILING DATE: 25-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-230921  
FILING DATE: 31-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-305605  
FILING DATE: 09-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/099,706  
FILING DATE: 30-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/769,996  
FILING DATE: 02-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/635,451  
FILING DATE: 28-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Mclelland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 900703G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930  
TELEFAX: (202) 887-0357  
INFORMATION FOR SEQ ID NO: 42:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 30 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-904-686A-42

Query Match 10.2%; Score 5; DB 2; Length 30;  
Best local Similarity 100.0%; Pred. No. 70;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 CDELA 33  
|||  
Db 16 CDELA 20

Search completed: December 4, 2002, 15:50:20  
Job time : 16 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:50:27 ; Search time 10 Seconds  
(Without alignments)  
79.587 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 49  
Sequence: 1 YLYQWLGAPVPPDPLKPRR.....DELADHTGFOEATRFYGPV 49

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 103943 seqs, 16242309 residues

Word size : 0

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCR\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/PCRUS\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	10	20.4	13	10	US-09-822-485-29
2	10	20.4	13	10	US-09-801-968-21
3	10	20.4	136	10	US-09-858-349-2
4	6	12.2	70	9	US-09-895-913A-364
5	6	12.2	70	10	US-09-815-242-11343
6	6	12.2	103	10	US-09-841-132-501
7	6	12.2	116	10	US-09-775-932-16
8	6	12.2	139	10	US-09-969-834-4
9	6	12.2	934	10	US-09-788-657-19
10	5	10.2	9	9	US-09-924-400-138
11	5	10.2	9	10	US-09-810-936-138
12	5	10.2	9	10	US-09-429-755-138
13	5	10.2	23	9	US-09-924-400-133
14	5	10.2	23	10	US-09-810-936-133
15	5	10.2	23	10	US-09-429-755-133
16	5	10.2	26	10	US-09-821-984-38
17	5	10.2	28	10	US-09-799-983-17
18	5	10.2	38	10	US-09-864-761-38378
19	5	10.2	40	10	US-09-925-299-1355

20	5	10.2	53	10	US-09-873-880-12	Sequence 12, Appl
21	5	10.2	57	10	US-09-864-761-36551	Sequence 36551, A
22	5	10.2	60	10	US-09-825-297-514	Sequence 514, App
23	5	10.2	63	10	US-09-867-550-1110	Sequence 1110, Ap
24	5	10.2	64	10	US-09-764-860-451	Sequence 451, App
25	5	10.2	67	10	US-09-864-761-45197	Sequence 45197, A
26	5	10.2	69	10	US-09-764-847-888	Sequence 888, App
27	5	10.2	74	10	US-09-925-302-509	Sequence 509, App
28	5	10.2	84	10	US-09-764-847-592	Sequence 592, App
29	5	10.2	86	10	US-09-864-761-33499	Sequence 33499, A
30	5	10.2	92	10	US-09-925-300-1795	Sequence 1795, Ap
31	5	10.2	94	10	US-09-764-877-1490	Sequence 1490, Ap
32	5	10.2	97	10	US-09-873-880-14	Sequence 14, Appl
33	5	10.2	99	10	US-09-864-761-40512	Sequence 40512, A
34	5	10.2	100	10	US-09-893-338-126	Sequence 126, App
35	5	10.2	102	10	US-09-758-308-2	Sequence 2, Appl1
36	5	10.2	105	9	US-09-984-245-182	Sequence 182, App
37	5	10.2	106	9	US-09-984-245-321	Sequence 321, App
38	5	10.2	109	10	US-09-864-761-36654	Sequence 36654, A
39	5	10.2	112	10	US-09-921-397-29	Sequence 29, Appl
40	5	10.2	119	10	US-09-925-301-1584	Sequence 1584, Ap
41	5	10.2	121	9	US-09-924-400-2	Sequence 2, Appl1
42	5	10.2	121	10	US-09-810-936-2	Sequence 2, Appl1
43	5	10.2	121	10	US-09-429-755-2	Sequence 2, Appl1
44	5	10.2	125	10	US-09-738-973-59	Sequence 59, Appl
45	5	10.2	133	10	US-09-853-918-61	Sequence 61, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-822-485-29
: Sequence 29, Application US/09822485
: Patent No. US20020001825A1
: GENERAL INFORMATION:
: APPLICANT: Itch, No. US20020001825A1uyuki
: TITLE OF INVENTION: NO. US20020001825A1 Fibroblast Growth Factor-Like Polypeptide
: FILE REFERENCE: 08035.0001-01000
: CURRENT APPLICATION NUMBER: US/09/822.485
: NUMBER OF SEQ. ID NOS: 35
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 29
: LENGTH: 13
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:protein tag
US-09-822-485-29

Query Match          20.4%; Score 10; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 GAPVPPDPL 16
Db      1 GAPVPPDPL 10

RESULT 2
US-09-801-968-21
: Sequence 21, Application US/09801968
: Patent No. US20020082205A1
: GENERAL INFORMATION:
: APPLICANT: Itch, No. US20020082205A1uyuki
: APPLICANT: Kavanaugh, W. Michael
: TITLE OF INVENTION: HUMAN RGF-23 GENE AND GENE EXPRESSION
: TITLE OF INVENTION: PRODUCTS
: FILE REFERENCE: PP-17150.001/201130.40901
: CURRENT APPLICATION NUMBER: US/09/801.968
: CURRENT FILING DATE: 2001-03-07
: NUMBER OF SEQ ID NOS: 46
```

SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: E tag  
US-09-801-968-21

Query Match  
Best Local Similarity 100.0%; Score 10; DB 10; Length 13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPDPL 16  
Db 1 GAVPYPDPL 10

RESULT 3  
US-09-858-349-2  
; Sequence 2, Application US/09858349  
; Patent No. US20020012909A1  
; GENERAL INFORMATION:  
; APPLICANT: PLAKSIN, Daniel  
; TITLE OF INVENTION: SMALL FUNCTIONAL UNITS OF ANTIBODY HEAVY CHAIN VARIABLE REGIONS  
; FILE REFERENCE: 87534-2800  
; CURRENT APPLICATION NUMBER: US/09/858,349  
; CURRENT FILING DATE: 2000-06-02  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: mouse hydridoma specific for H-2D + RGPGRATVTI peptide  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (99)..(107)  
; OTHER INFORMATION: variable  
US-09-858-349-2

Query Match  
Best Local Similarity 100.0%; Score 10; DB 10; Length 136;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPDPL 16  
Db 122 GAVPYPDPL 131

RESULT 4  
US-09-895-913A-364  
; Sequence 364, Application US/09895913A  
; Patent No. US20020160456A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleanthous, Harold  
; APPLICANT: Al-Garawi, Amal  
; APPLICANT: Miller, Charles  
; APPLICANT: Tomb, Jean Francois  
; APPLICANT: Oomen, Raymond P.  
; TITLE OF INVENTION: Identification of Polynucleotides  
; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in the  
; FILE REFERENCE: 06132/043002  
; CURRENT APPLICATION NUMBER: US/09/895,913A  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 08/881,227  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 364  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori

US-09-895-913A-364  
Query Match  
Best Local Similarity 100.0%; Score 6; DB 9; Length 70;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EAYRRF 45  
Db 14 EAYRRF 19

RESULT 5  
US-09-815-242-11343  
; Sequence 11343, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11343  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-815-242-11343

Query Match  
Best Local Similarity 100.0%; Score 6; DB 10; Length 70;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EAYRRF 45  
Db 14 EAYRRF 19

RESULT 6  
US-09-841-132-501  
; Sequence 501, Application US/09841132  
; Patent No. US20020061848A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C8  
; CURRENT APPLICATION NUMBER: US/09/841,132  
; CURRENT FILING DATE: 2001-04-23



```
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 501
; LENGTH: 103
; TYPE: PRF
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-501

Query Match
Best Local Similarity 12.2%; Score 6; DB 10; Length 103;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLXQWL 6,
Db 20 YLXQWL 25

RESULT 7
US-09-775-932-16
; Sequence 16, Application US/09775932
; Patent No. US20020137671A1
; GENERAL INFORMATION:
; APPLICANT: University of British Columbia
; TITLE OF INVENTION: Production and use of Modified Cystatins
; FILE REFERENCE: 58069
; CURRENT APPLICATION NUMBER: US/09/775,932
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: CA99/00717
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/095,503
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 116
; TYPE: PRF
; ORGANISM: Gallus sp.
US-09-775-932-16

Query Match
Best Local Similarity 12.2%; Score 6; DB 10; Length 116;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11
Db 8 LGAPVP 13

RESULT 8
US-09-969-834-4
; Sequence 4, Application US/09969834
; Patent No. US20020102711A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; OPERATING SYSTEM: IBM compatible
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/969,834
; FILING DATE: 01-Oct-2001
```

```
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/471,765
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/791,522
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 09/471,765
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 118195
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-969-834-4

Query Match
Best Local Similarity 12.2%; Score 6; DB 10; Length 139;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11
Db 31 LGAPVP 36

RESULT 9
US-09-788-657-19
; Sequence 19, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaidis, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; FILE REFERENCE: 01107,00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 934
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-788-657-19

Query Match
Best Local Similarity 12.2%; Score 6; DB 10; Length 934;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12
Db 614 GAPVPY 619

RESULT 10
US-09-924-400-138
; Sequence 138, Application US/09924400
```

```
; Patent No. US20020165371A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Lynda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.419C12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted HLA A2.1 Motifs (T-cell epitopes)
US-09-924-400-138
```

```
Query Match      10.2%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      39 QEAYR 43
DB      3 QEAYR 7
```

```
RESULT 11
US-09-810-936-138
; Sequence 138, Application US/09810936
; Patent No. US20020068285A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Linda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.419C11
; CURRENT APPLICATION NUMBER: US/09/810,936
; CURRENT FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 138
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted HLA A2.1 Motifs (T-cell epitopes)
US-09-810-936-138
```

```
Query Match      10.2%; Score 5; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      39 QEAYR 43
```

```
DB      3 QEAYR 7
RESULT 12
US-09-429-755-138
; Sequence 138, Application US/09429755A
; Patent No. US2002011467A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.419C6
; CURRENT APPLICATION NUMBER: US/09/429,755A
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 138
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted HLA A2.1 Motifs (T-cell epitopes)
US-09-429-755-138
```

```
Query Match      10.2%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      39 QEAYR 43
DB      3 QEAYR 7
```

```
RESULT 13
US-09-924-400-133
; Sequence 133, Application US/09924400
; Patent No. US20020165371A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Lynda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.419C12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted Th Motifs (B-cell epitopes)
US-09-924-400-133
```

```
Query Match      10.2%; Score 5; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 24;
```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 QEAYR 43  
|1111|  
Db 10 QEAYR 14

## RESULT 14

US-09-810-936-133  
; Sequence 133, Application US/09810936  
; Patent No. US20020068285A1  
; GENERAL INFORMATION:  
; APPLICANT: Fridakis, Tony N.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Smith, John M.  
; APPLICANT: Misher, Linda E.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Wang, Aljun  
; APPLICANT: Skelky, Yasir A.W.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Day, Craig H.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C11  
; CURRENT APPLICATION NUMBER: US/09/810,936  
; CURRENT FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 334  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 133  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Predicted Th Motifs (B-cell epitopes)  
US-09-810-936-133

Query Match 10.2%; Score 5; DB 10; Length 23;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
|1111|  
Db 10 QEAYR 14

## RESULT 15

US-09-429-755-133  
; Sequence 133, Application US/09429755A  
; Patent No. US2002011467A1  
; GENERAL INFORMATION:  
; APPLICANT: Fridakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Misher, Lynda  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C6  
; CURRENT APPLICATION NUMBER: US/09/429,755A  
; CURRENT FILING DATE: 1999-10-28  
; NUMBER OF SEQ ID NOS: 315  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 133  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Predicted Th Motifs (B-cell epitopes)  
US-09-429-755-133

Query Match 10.2%; Score 5; DB 10; Length 23;

Best Local Similarity 100.0%; Pred. No. 24;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
|1111|  
Db 10 QEAYR 14

## RESULT 16

US-09-821-984-38  
; Sequence 38, Application US/09821984  
; Patent No. US2002004205A1  
; GENERAL INFORMATION:  
; APPLICANT: Consler, Thomas G.  
; APPLICANT: Iannone, Marie A.  
; APPLICANT: Gray, John G.  
; APPLICANT: Stimmel, Julia E.  
; TITLE OF INVENTION: METHOD OF INVESTIGATING FUNCTIONAL  
; TITLE OF INVENTION: MOLECULAR INTERACTIONS AND REAGENTS FOR USE THEREIN  
; FILE REFERENCE: 07083.000702  
; CURRENT APPLICATION NUMBER: US/09/821,984  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 60/193,826  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 38  
; LENGTH: 26  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/note -  
; OTHER INFORMATION: synthetic construct  
US-09-821-984-38

Query Match 10.2%; Score 5; DB 10; Length 26;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LADHI 36  
|1111|  
Db 12 LADHI 16

## RESULT 17

US-09-799-983-17  
; Sequence 17, Application US/09799983  
; Patent No. US2001002903A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaml, Paul  
; APPLICANT: Parker, Charles  
; TITLE OF INVENTION: NOVEL GENE RNO UPREGULATED IN LEUKEMIA CELLS BY NITRIC OXIDE  
; FILE REFERENCE: 1321.2.51  
; CURRENT APPLICATION NUMBER: US/09/799,983  
; CURRENT FILING DATE: 2001-03-06  
; PRIOR APPLICATION NUMBER: US 60/186,971  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-983-17

Query Match 10.2%; Score 5; DB 10; Length 28;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELA 33  
|1111|  
Db 17 CDELA 21

RESULT 18  
US-09-864-761-38378  
; Sequence 38378, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263,6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 38378  
; LENGTH: 38  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005630.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.2  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.5  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.9  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.7  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.7  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.3  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.5  
; OTHER INFORMATION: EST\_HUMAN HIT: AW162304.1, EVALUOE 3.00e-05  
US-09-864-761-38378

Query Match 10.2%; Score 5; DB 10; Length 38;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVP 11  
Db 24 GAPVP 28

RESULT 19  
US-09-925-299-1355  
; Sequence 1355, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1355  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (25)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (34)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (36)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1355

Query Match 10.2%; Score 5; DB 10; Length 40;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVP 11  
Db 28 GAPVP 32

RESULT 20  
US-09-873-880-12  
; Sequence 12, Application US/09873880  
; Patent No. US20020123118A1  
; GENERAL INFORMATION:  
; APPLICANT: Sewalt, Vincent  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: GLYCINE METABOLISM ENZYMES  
; FILE REFERENCE: BB1192 US CIP  
; CURRENT APPLICATION NUMBER: US/09/873,880  
; CURRENT FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 09/363,321  
; PRIOR FILING DATE: July 28, 1999  
; PRIOR APPLICATION NUMBER: 60/094,839  
; PRIOR FILING DATE: July 31, 1998  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Microsoft Office 97  
SEQ ID NO 12  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-873-880-12

Query Match 10.2%; Score 5; DB 10; Length 53;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 30 DELAD 34  
11111  
Db 46 DELAD 50

RESULT 21  
US-09-864-761-36551  
Sequence 36551, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aemica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 36551  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AP000247.1  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.67  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.74  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.75  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.85  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.65  
OTHER INFORMATION: EST HUMAN HIT: BE968985.1, EVALUATE 7.20e-01  
OTHER INFORMATION: SWISSPROT HIT: Q13009, EVALUATE 1.00e-31  
US-09-864-761-36551

Query Match 10.2%; Score 5; DB 10; Length 57;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 NPDCD 30  
11111  
Db 14 NPDCD 18

RESULT 22  
US-09-925-297-514  
Sequence 514, Application US/09925297  
Patent No. US20020081659A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA105  
CURRENT APPLICATION NUMBER: US/09/925,297  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05989  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 928  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 514  
LENGTH: 60  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (1)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (6)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (7)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (8)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-297-514

Query Match 10.2%; Score 5; DB 10; Length 60;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 FYGPV 49  
11111  
Db 35 FYGPV 39

RESULT 23  
US-09-867-550-1110  
Sequence 1110, Application US/09867550  
Patent No. US20020082206A1  
GENERAL INFORMATION:  
APPLICANT: Leach, Martin D.  
APPLICANT: Mehraban, Foad  
APPLICANT: Conley, Pamela  
APPLICANT: Law, Debbie  
APPLICANT: Topper, James  
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a  
FILE REFERENCE: 21402-013 (Cura-313)  
CURRENT APPLICATION NUMBER: US/09/867,550

```
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1110
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1110

Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 YDPL 16
DB 29 YDPL 33

RESULT 24
US-09-764-860-451
; Sequence 451, Application US/09764860
; Patent No. US2002009493A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA REMOVED - consult PALM or file wrapper
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 451
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-860-451

Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 YOMLG 7
DB 51 YOMLG 55

RESULT 25
US-09-864-761-45197
; Sequence 45197, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-x-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
```

```
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263..6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45197
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121673..27
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.83
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.81
; OTHER INFORMATION: SWISSPROT HIT: P26807, EVALUATE 2.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: AW438701.1, EVALUATE 1.90e+00
US-09-864-761-45197

Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVP 11
DB 31 GAPVP 35

RESULT 26
US-09-764-847-888
; Sequence 888, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/764,847
; PRIOR APPLICATION DATA REMOVED - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 888
; LENGTH: 69
```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-847-888

Query Match  
Best Local Similarity 10.2%; Score 5; DB 10; Length 69;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WIGAP 9  
|||||  
DB 47 WIGAP 51

RESULT 27  
US-09-925-302-509

Sequence 509, Application US/09925302  
Patent No. US20020044941A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA104  
CURRENT APPLICATION NUMBER: US/09/925,302  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05918  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 896  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 509  
LENGTH: 74  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (31)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-302-509

Query Match  
Best Local Similarity 10.2%; Score 5; DB 10; Length 74;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 HIGFO 39  
|||||  
DB 4 HIGFO 8

RESULT 28  
US-09-764-847-592

Sequence 592, Application US/09764847  
Patent No. US20020132767A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC009  
CURRENT APPLICATION NUMBER: US/09/764,847  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 2003  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 592  
LENGTH: 84  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (41)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (61)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-847-592

Query Match  
Best Local Similarity 10.2%; Score 5; DB 10; Length 84;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPV 10  
|||||  
DB 65 LGAPV 69

RESULT 29  
US-09-864-761-33499

Sequence 33499, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 33499  
LENGTH: 86  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL035603.11  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 32  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.9  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 27  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 20  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 50

```
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 27
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.5
; OTHER INFORMATION: SWISSPROT HIT: P79103, EVALUATE 9.00e-30
; OTHER INFORMATION: EST_HUMAN HIT: W00563.1, EVALUATE 5.00e-34
US-09-864-761-33499

Query Match          10.2%; Score 5; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 YPDP 16
    |||||
Db 53 YPDP 57

RESULT 30
US-09-925-300-1795
; Sequence 1795, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruden,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1795
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1795

Query Match          10.2%; Score 5; DB 10; Length 92;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 YPDP 16
    |||||
Db 53 YPDP 57

RESULT 31
US-09-764-877-1490
; Sequence 1490, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PAM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1490
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (94)
```

```
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1490

Query Match          10.2%; Score 5; DB 10; Length 94;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADH 35
    |||||
Db 68 ELADH 72

RESULT 32
US-09-873-880-14
; Sequence 14, Application US/09873880
; Patent No. US20020123118A1
; GENERAL INFORMATION:
; APPLICANT: Sewalt, Vincent
; APPLICANT: Falco, S. Carl
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: GLYCINE METABOLISM ENZYMES
; FILE REFERENCE: BB1192 US CIP
; CURRENT APPLICATION NUMBER: US/09/873,880
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/363,321
; PRIOR FILING DATE: July 28, 1999
; PRIOR APPLICATION NUMBER: 60/094,839
; PRIOR FILING DATE: July 31, 1998
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-873-880-14

Query Match          10.2%; Score 5; DB 10; Length 97;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 DELAD 34
    |||||
Db 46 DELAD 50

RESULT 33
US-09-864-761-40512
; Sequence 40512, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```



```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-893-238-126
Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 30 DELAD 34
Db 37 DELAD 41
RESULT 34
US-09-893-238-126
; Sequence 126, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 126
; LENGTH: 100
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-893-238-126
Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 100;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 EAYRR 44
Db 63 EAYRR 67
RESULT 35
US-09-758-308-2
; Sequence 2, Application US/09758308
; Patent No. US20020090607A1
; GENERAL INFORMATION:
; APPLICANT: HOWARD A. FIELDS AND YURY E. KHUDYAKOV
; TITLE OF INVENTION: ANTIGENIC EPITOPES AND MOSAIC POLYPEPTIDES OF HEPATITIS C VIR
; FILE REFERENCE: 14114.034902
; CURRENT APPLICATION NUMBER: US/09/758,308
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/092,339
; PRIOR FILING DATE: 1999-07-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
; US-09-758-308-2
Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 102;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 CDELA 33
Db 44 CDELA 48
RESULT 36
US-09-984-245-182
; Sequence 182, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
```

;; PRIOR APPLICATION NUMBER: US 60/050,937  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,187  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,099  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,352  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,186  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,069  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,095  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,131  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,096  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,355  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,160  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,351  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,154  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/054,804  
;; PRIOR FILING DATE: 1997-08-05  
;; PRIOR APPLICATION NUMBER: US 60/056,370  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: US 60/060,862  
;; PRIOR FILING DATE: 1997-10-02  
;; NUMBER OF SEQ ID NOS: 343  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 182  
;; LENGTH: 105  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-984-245-182

Query Match 10.2% Score 5; DB 9; Length 105;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 PYDP 15  
Db 85 PYDP 89  
RESULT 37  
US-09-984-245-321  
;; Sequence 321, Application US/09984245  
;; Patent No. US20020165374A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Young et al.  
;; TITLE OF INVENTION: 87 Human Secreted Proteins  
;; FILE REFERENCE: P2004P1  
;; CURRENT APPLICATION NUMBER: US/09/984,245  
;; CURRENT FILING DATE: 2001-10-29  
;; PRIOR APPLICATION NUMBER: 09/154,707  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: PCT/US98/05311  
;; PRIOR FILING DATE: 1998-03-19  
;; PRIOR APPLICATION NUMBER: US 60/041,277  
;; PRIOR FILING DATE: 1997-03-21  
;; PRIOR APPLICATION NUMBER: US 60/042,344  
;; PRIOR FILING DATE: 1997-03-21  
;; PRIOR APPLICATION NUMBER: US 60/041,276  
;; PRIOR FILING DATE: 1997-03-21  
;; PRIOR APPLICATION NUMBER: US 60/041,281  
;; PRIOR FILING DATE: 1997-03-21  
;; PRIOR APPLICATION NUMBER: US 60/048,094  
;; PRIOR FILING DATE: 1997-05-30

;; PRIOR APPLICATION NUMBER: US 60/048,350  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,188  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,135  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/050,937  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,187  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,099  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,352  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,186  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,069  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,095  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,131  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,096  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,351  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,154  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/054,804  
;; PRIOR FILING DATE: 1997-08-05  
;; PRIOR APPLICATION NUMBER: US 60/056,370  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: US 60/060,862  
;; PRIOR FILING DATE: 1997-10-02  
;; NUMBER OF SEQ ID NOS: 343  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 321  
;; LENGTH: 106  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-984-245-321

Query Match 10.2% Score 5; DB 9; Length 106;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 PYDP 15  
Db 86 PYDP 90  
RESULT 38  
US-09-864-761-36654  
;; Sequence 36654, Application US/09864761  
;; Patent No. US20020048763A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Penn, Sharon G.  
;; APPLICANT: Rank, David R.  
;; APPLICANT: Hanzel, David K.  
;; APPLICANT: Chen, Wensheng  
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
;; FILE REFERENCE: Aecm1ca-X-1  
;; CURRENT APPLICATION NUMBER: US/09/864,761  
;; CURRENT FILING DATE: 2001-05-23  
;; PRIOR APPLICATION NUMBER: US 60/180,312  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: US 60/207,456  
;; PRIOR FILING DATE: 2000-05-26

```
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36654
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049553.15
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.1
; OTHER INFORMATION: EST_HUMAN HIT: A1924653.1, EVALUATE 3.00e-58
; OTHER INFORMATION: SWISSPROT HIT: Q62718, EVALUATE 2.00e-04
US-09-864-761-36654

Query Match          10.2%: Score 5; DB 10; Length 109;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 7 GAPVP 11
    |||||
Db 93 GAPVP 97

RESULT 39
US-09-921-397-29
; Sequence 29, Application US/09921397
; Patent No. US20020151484A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIDENICS
; TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
; TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
; FILE REFERENCE: B4809A - JAZ
```

```
; CURRENT APPLICATION NUMBER: US/09/921.397
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: EP 00402225.7
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-29
```

```
Query Match          10.2%: Score 5; DB 10; Length 112;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 29 CDELA 33
    |||||
Db 30 CDELA 34
```

```
RESULT 40
US-09-925-301-1584
; Sequence 1584, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1584
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (60)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (99)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (118)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1584

Query Match          10.2%: Score 5; DB 10; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 YPDPL 16
    |||||
Db 84 YPDPL 88

RESULT 41
US-09-924-400-2
; Sequence 2, Application US/09924400
; Patent No. US20020165371A1
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Lynda E.
; APPLICANT: Dillon, David C.
; APPLICANT: Retter, Marc W.
```

```

; APPLICANT: Wang, Aljun
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419c12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-400-2
```

```

Query Match          10.2%; Score 5; DB 9; Length 121;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 39 QEAYR 43
      |||||
DB 63 QEAYR 67
```

```

RESULT 42
US-09-810-936-2
; Sequence 2, Application US/09810936
; Patent No. US20020068285A1
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Linda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aljun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419c11
; CURRENT APPLICATION NUMBER: US/09/810,936
; CURRENT FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-810-936-2
```

```

Query Match          10.2%; Score 5; DB 10; Length 121;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 39 QEAYR 43
      |||||
DB 63 QEAYR 67
```

```

RESULT 43
US-09-429-755-2
; Sequence 2, Application US/09429755A
; Patent No. US20020111467A1
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
```

```

; APPLICANT: Misher, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419c6
; CURRENT APPLICATION NUMBER: US/09/429,755A
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-429-755-2
```

```

Query Match          10.2%; Score 5; DB 10; Length 121;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 39 QEAYR 43
      |||||
DB 63 QEAYR 67
```

```

RESULT 44
US-09-738-973-59
; Sequence 59, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475c9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-59
```

```

Query Match          10.2%; Score 5; DB 10; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 11 PYDP 15
      |||||
DB 37 PYDP 41
```

```

RESULT 45
US-09-853-918-61
; Sequence 61, Application US/09853918
; Patent No. US20020068346A1
; GENERAL INFORMATION:
; APPLICANT: Krystek, Stanley R.
; APPLICANT: Sherliff, Steven
; APPLICANT: Wilmer, Mark R.
; APPLICANT: Hollenbaugh, Diane L.
```

; APPLICANT: Yan, Ning  
; APPLICANT: Mouravieff, Julie E.  
; APPLICANT: Einspahr, Howard M.  
; APPLICANT: Kish, Kevin  
; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE  
; FILE REFERENCE: DB24NP  
; CURRENT APPLICATION NUMBER: US/09/853,918  
; PRIOR FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: 60/203,448  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 61  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-918-61

Query Match 10.2%; Score 5; DB 10; Length 133;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 DCDL 32  
|||  
Db 104 DCDL 108

Search completed: December 4, 2002, 15:54:27  
Job time : 10 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:44:21 : Search time 14 seconds  
(without alignments)  
102.980 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 272  
Sequence: 1 YLYQWLGAPVPPDPLXPRR.....DELADHIGFQEARFRFGPV 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents-AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCUTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfillsl.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266	97.8	49	6	Patent No. 5164483-1
2	266	97.8	49	6	Patent No. 5164483-2
3	266	97.8	49	6	Patent No. 5434245-1
4	266	97.8	50	6	Patent No. 5434245-2
5	266	97.8	50	6	Patent No. 5434245-3
6	266	97.8	51	6	Patent No. 5434245-4
7	266	97.8	98	6	Patent No. 5164483-3
8	242.5	89.2	48	6	Patent No. 5168041-1
9	70	25.7	13	2	US-08-796-598-13
10	70	25.7	13	2	US-08-447-175A-13
11	70	25.7	13	2	US-08-943-915-6
12	70	25.7	13	3	US-08-881-037-112
13	70	25.7	13	3	US-08-652-816A-43
14	62	22.8	13	4	US-09-142-974B-5
15	57.5	21.1	423	1	US-08-523-376-3
16	54	19.9	106	4	US-09-083-351-7
17	54	19.9	106	4	US-09-083-352-7
18	53.5	19.7	486	1	US-07-672-483-2
19	53	19.5	431	4	US-09-038-832-2
20	53	19.5	431	4	US-09-038-832-4
21	53	19.5	2254	2	US-08-797-010-3
22	53	19.5	2254	2	US-08-797-010-3
23	52.5	19.3	12	3	US-08-968-747-7
24	52.5	19.3	12	3	US-08-993-071-28
25	52.5	19.3	486	4	US-08-259-451-13
26	52	19.1	11	6	Patent No. 5168041-2
27	51	18.8	334	1	US-08-482-385A-7

28	50.5	18.6	477	1	US-08-136-922-2	Sequence 2, Appl1
29	50.5	18.6	771	1	US-08-121-713D-54	Sequence 54, Appl
30	50.5	18.6	771	1	US-08-835-268-54	Sequence 54, Appl
31	50.5	18.6	771	2	US-09-060-692-54	Sequence 54, Appl
32	50.5	18.6	771	3	US-08-833-391-54	Sequence 54, Appl
33	50.5	18.6	771	4	US-09-060-610-54	Sequence 54, Appl
34	50.5	18.6	771	5	PCR-US94-10151A-54	Sequence 54, Appl
35	50.5	18.6	1336	2	US-08-231-193A-58	Sequence 58, Appl
36	50.5	18.6	1336	2	US-08-486-273A-58	Sequence 58, Appl
37	50.5	18.6	1336	3	US-08-940-086A-58	Sequence 58, Appl
38	50.5	18.6	1336	4	US-08-940-035A-58	Sequence 58, Appl
39	50.5	18.6	1336	4	US-08-935-105A-58	Sequence 58, Appl
40	50.5	18.6	1336	4	US-09-648-797-58	Sequence 58, Appl
41	49.5	18.2	393	3	US-08-979-917A-3	Sequence 3, Appl1
42	49.5	18.2	393	4	US-09-153-599A-11	Sequence 11, Appl
43	49.5	18.2	2396	1	US-08-157-005-2	Sequence 2, Appl1
44	49.5	18.2	2396	1	US-08-747-863-2	Sequence 2, Appl1
45	49.5	18.2	2396	4	US-09-565-86A-2	Sequence 2, Appl1

## ALIGNMENTS

RESULT 1  
5164483-1  
; Patent No. 5164483  
; APPLICANT: Takashi, Kurihara;Eiji, Taniyama;Sachio, Hirose  
; TITLE OF INVENTION: Y-CARBOXYGLUTAMATE DERIVATIVE METHOD  
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN  
; USING THE SAME  
; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US-07/575, 639  
; FILING DATE: 31-AUG-1990  
; SEQ ID NO: 1:  
; LENGTH: 49  
5164483-1

Query Match 97.8%; Score 266; DB 6; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.8e-30;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPPDPLXPRRXVCXLNPDDELADHIGFQEARFRFGPV 49  
DB 1 YLYQWLGAPVPPDPLXPRRXVCXLNPDDELADHIGFQEARFRFGPV 49

RESULT 2  
5164483-2  
; Patent No. 5164483  
; APPLICANT: Takashi, Kurihara;Eiji, Taniyama;Sachio, Hirose  
; TITLE OF INVENTION: Y-CARBOXYGLUTAMATE DERIVATIVE METHOD  
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN  
; USING THE SAME  
; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US-07/575, 639  
; FILING DATE: 31-AUG-1990  
; SEQ ID NO: 2:  
; LENGTH: 49  
5164483-2

Query Match 97.8%; Score 266; DB 6; Length 49;  
Best Local Similarity 98.0%; Pred. No. 1.8e-30;  
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPPDPLXPRRXVCXLNPDDELADHIGFQEARFRFGPV 49  
DB 1 YLYQWLGAPVPPDPLXPRRXVCXLNPDDELADHIGFQEARFRFGPV 49

RESULT 3  
5434245-1

; Patent No. 5434245  
; APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/993,980  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989  
; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989  
; SEQ ID NO: 1:  
; LENGTH: 49  
5434245-1

Query Match 97.8%; Score 266; DB 6; Length 49;  
Best Local Similarity 93.9%; Pred. No. 1.8e-30;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 49  
DB 1 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 49

RESULT 4  
5434245-2  
; Patent No. 5434245  
; APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/993,980  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989  
; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989  
; SEQ ID NO: 2:  
; LENGTH: 50  
5434245-2

Query Match 97.8%; Score 266; DB 6; Length 50;  
Best Local Similarity 93.9%; Pred. No. 1.8e-30;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 49  
DB 1 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 49

RESULT 5  
5434245-3  
; Patent No. 5434245  
; APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/993,980  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989  
; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989  
; SEQ ID NO: 3:  
; LENGTH: 50  
5434245-3

Query Match 97.8%; Score 266; DB 6; Length 50;  
Best Local Similarity 93.9%; Pred. No. 1.8e-30;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 49  
DB 2 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 50

RESULT 6  
5434245-4  
; Patent No. 5434245  
; APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/993,980  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989  
; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989  
; SEQ ID NO: 4:  
; LENGTH: 51  
5434245-4

Query Match 97.8%; Score 266; DB 6; Length 51;  
Best Local Similarity 93.9%; Pred. No. 1.9e-30;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 49  
DB 2 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 50

RESULT 7  
5164483-3  
; Patent No. 5164483  
; APPLICANT: Takashi, Kurihara; Eiji, Taniyama; Sachio, Hirose  
; TITLE OF INVENTION: Y-CARBOXYGLUTAMATE DERIVATIVE METHOD  
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN  
; USING THE SAME  
; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/575,639  
; FILING DATE: 31-AUG-1990  
; SEQ ID NO: 3:  
; LENGTH: 110  
5164483-3

Query Match 97.8%; Score 266; DB 6; Length 98;  
Best Local Similarity 100.0%; Pred. No. 4e-30;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 49  
DB 1 YLYQWLGAPVPPDPLXPRRYXCXINPDCDELADHIGFQEAAYRRFGPV 49

RESULT 8  
5168041-1  
; Patent No. 5168041  
; APPLICANT: BERGMANN, ANDREAS E.  
; TITLE OF INVENTION: METHOD FOR THE DETERMINATION OF  
; OSTEOCALCIN IN HUMAN SERUM OR PLASMA  
; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/416,728  
; FILING DATE: 03-OCT-1989  
; SEQ ID NO: 1:  
; LENGTH: 48  
5168041-1





REGISTRATION NUMBER: 27,657  
REFERENCE/DOCKET NUMBER: A-469  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805.447.4112  
TELEFAX: 805.447.1090  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-943-915-6

Query Match 25.7%; Score 70; DB 2; Length 13;  
Best Local Similarity 92.3%; Pred. No. 0.0012;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 GAVPYPPDLXPR 19  
|||||  
Db 1 GAVPYPPDLXPR 13

RESULT 12  
US-08-861-037-112  
Sequence 112, Application US/08881037  
Patent No. 6080588  
GENERAL INFORMATION:  
APPLICANT: Glick, Gary D.  
APPLICANT: Swanson, Patrick C.  
TITLE OF INVENTION: DNA BINDING ANTIBODIES  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/881,037  
FILING DATE: 23-JUN-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/443,540  
FILING DATE: 18-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Konaki, Antoinette F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: 203442110710  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 813-5600  
TELEFAX: (650) 494-0792  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-861-037-112

Query Match 25.7%; Score 70; DB 3; Length 13;  
Best Local Similarity 92.3%; Pred. No. 0.0012;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 GAVPYPPDLXPR 19

|||||  
Db 1 GAVPYPPDLXPR 13

RESULT 13  
US-08-652-816A-43  
Sequence 43, Application US/08652816A  
Patent No. 5872215  
GENERAL INFORMATION:  
APPLICANT: Osbourn, JK  
APPLICANT: Allen, DJ  
APPLICANT: McCafferty, JG  
TITLE OF INVENTION: Specific binding members, materials and  
NUMBER OF SEQUENCES: 53  
METHODS:  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,816A  
FILING DATE: 23-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125579.4  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125579.8  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206318.9  
FILING DATE: 24-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206372.6  
FILING DATE: 23-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9525004.9  
FILING DATE: 07-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610824.6  
FILING DATE: 23-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/02240  
FILING DATE: 02-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/244,597  
FILING DATE: 01-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/33308  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-652-816A-43

Query Match 25.7%; Score 70; DB 2; Length 43;  
Best Local Similarity 92.3%; Pred. No. 0.0049;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 GAVPYPPDLXPR 19  
|||||

Db 23 GAPVPYDPLEPR 35

## RESULT 14

US-09-142-974B-5

; Sequence 5, Application US/09142974B

; Patent No. 6451995

## ; GENERAL INFORMATION:

; APPLICANT: Cheung, Nai-Kong V.

; APPLICANT: Larson, Steven M.

; APPLICANT: Guo, Hong-Fen

; APPLICANT: Rivlin, Ken

; APPLICANT: Sadelaïn, Michel

; TITLE OF INVENTION: Single Chain Fv Constructs of Anti-ganglioside GD2

; FILE REFERENCE: MSK.P-013-USNP

; CURRENT APPLICATION NUMBER: US/09/142,974B

; CURRENT FILING DATE: 1998-09-18

; PRIOR APPLICATION NUMBER: PCT/US97/04427

; PRIOR FILING DATE: 1997-03-20

; PRIOR APPLICATION NUMBER: 60/013,703

; PRIOR FILING DATE: 1996-03-20

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 5

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: antibody tag

US-09-142-974B-5

Query Match 22.8%; Score 62; DB 4; Length 13;

Best Local Similarity 84.6%; Pred. No. 0.016;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 7 GAPVPYDPLXPR 19

Db 1 GAPVPYDPLEPR 13

## RESULT 15

US-08-523-376-3

; Sequence 3, Application US/08523376

; Patent No. 5808030

## ; GENERAL INFORMATION:

; APPLICANT: Tsutomu, FUJIMURA

; APPLICANT: Satoshi, TAKEDA

; APPLICANT: Yoshikazu, SHIMADA

; APPLICANT: Kouichi, Ozaki

; APPLICANT: Sadahito, SIN

; TITLE OF INVENTION: htfliiA GENE

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak &amp; Seas

; STREET: 2100 Pennsylvania Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: United States

; ZIP: 20037-3202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentln Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/523,376

; FILING DATE:

; CLASSIFICATION: 536

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 293-7060

; TELEFAX: (202) 293-7860

; TELEX: 6491103

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 423 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-523-376-3

Query Match 21.1%; Score 57.5; DB 1; Length 423;

Best Local Similarity 36.6%; Pred. No. 4;

Matches 15; Conservative 1; Mismatches 12; Indels 13; Gaps 3;

OY 8 APVPYDPLXPRRXVCXINPDCD-----ELADHTG 37

Db 86 APVP-PRPALPRRIFCSF-PDCSANTSKAWKLDHILCKHTG 124

Search completed: December 4, 2002, 15:48:22

Job time: 14 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:45:01 ; Search time 11 Seconds

(without alignments)  
72.352 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 272

Sequence: 1 YLYQWLGAPVPPDPLXPRR.....DELAHIGQEARRYGV 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA:\*

1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	70	25.7	13	10	US-09-822-485-29
2	70	25.7	13	10	US-09-801-968-21
3	70	25.7	136	10	US-09-858-349-2
4	59.5	21.9	677	10	US-09-815-242-10210
5	55	20.2	340	10	US-09-814-777A-15
6	55	20.2	384	10	US-09-814-777A-18
7	55	20.2	384	10	US-09-814-777A-20
8	55	20.2	470	10	US-09-814-777A-100
9	54.5	20.0	677	10	US-09-815-242-11921
10	54.5	20.0	704	10	US-09-815-242-11925
11	54	19.9	216	10	US-09-924-358-47
12	53	19.5	468	10	US-09-814-777A-2
13	53	19.5	468	10	US-09-814-777A-4
14	53	19.5	499	10	US-09-972-714-10
15	52	19.1	422	10	US-09-765-068-2
16	52	19.1	459	10	US-09-797-033-5
17	51	18.8	191	9	US-09-950-933A-66
18	51	18.8	565	9	US-09-999-248-4
19	50.5	18.6	1336	9	US-09-945-901-58

20	50.5	18.6	1336	9	US-10-007-747-58	Sequence 58, Appl
21	50	18.4	103	10	US-09-841-132-501	Sequence 50, Appl
22	50	18.4	271	10	US-09-846-808-20	Sequence 20, Appl
23	48	17.6	162	10	US-09-925-299-1151	Sequence 1151, Ap
24	48	17.6	214	10	US-09-833-503A-4	Sequence 4, Appl
25	48	17.6	224	10	US-09-925-301-1018	Sequence 1018, Ap
26	47	17.3	43	9	US-10-041-406-6	Sequence 6, Appl
27	47	17.3	409	10	US-09-815-242-11699	Sequence 11699, A
28	47	17.3	412	10	US-09-795-693-21	Sequence 21, Appl
29	47	17.3	632	9	US-09-981-353-50	Sequence 50, Appl
30	47	17.3	907	9	US-10-008-739A-2	Sequence 2, Appl
31	46.5	17.1	198	10	US-09-799-848-5	Sequence 1380, Ap
32	46.5	17.1	219	10	US-09-925-301-1380	Sequence 888, App
33	46.5	17.1	219	10	US-09-764-864-1338	Sequence 1338, Ap
34	46.5	17.1	219	10	US-09-764-864-1338	Sequence 8, Appl
35	46	16.9	213	10	US-09-764-864-1154	Sequence 2, Appl
36	46	16.9	879	10	US-09-799-875-2	Sequence 2, Appl
37	46	16.9	1550	10	US-09-995-542-6	Sequence 6, Appl
38	46	16.9	2100	10	US-09-995-542-6	Sequence 5, Appl
39	46	16.9	2144	10	US-09-858-194-2	Sequence 10, Appl
40	46	16.9	2146	10	US-09-995-542-5	Sequence 5, Appl
41	46	16.9	2257	10	US-09-767-479-10	Sequence 18, Appl
42	45.5	16.7	189	9	US-09-905-291A-18	Sequence 18, Appl
43	45.5	16.7	189	10	US-09-909-320-18	Sequence 18, Appl
44	45.5	16.7	189	10	US-09-909-088B-18	Sequence 571, App
45	45	16.5	104	10	US-09-841-132-571	

#### ALIGNMENTS

RESULT 1  
US-09-822-485-29  
; Sequence 29, Application US/09822485  
; Patent No. US20020001825A1  
; GENERAL INFORMATION:  
; APPLICANT: Itch, No. US0020001825A1uyuk1  
; TITLE OF INVENTION: No. US0020001825A1el Fibroblast Growth Factor-Like Polypeptide  
; FILE REFERENCE: 08035.0001-01000  
; CURRENT APPLICATION NUMBER: US/09/822,485  
; CURRENT FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:protein tag  
US-09-822-485-29

Query Match 25.7% Score 70: DB 10: Length 13:  
Best Local Similarity 92.3% Pred. No. 0.00085:  
Matches 12: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

QY 7 GAPVPPDPLXPR 19  
Db 1 GAPVPPDPLXPR 13

RESULT 2  
US-09-801-968-21  
; Sequence 21, Application US/09801968  
; Patent No. US20020082205A1  
; GENERAL INFORMATION:  
; APPLICANT: Itch, No. US20020082205A1uyuk1  
; APPLICANT: Kavanaugh, W. Michael  
; TITLE OF INVENTION: HUMAN GGF-23 GENE AND GENE EXPRESSION  
; TITLE OF INVENTION: PRODUCTS  
; FILE REFERENCE: PP-17150.001/201130.40901  
; CURRENT APPLICATION NUMBER: US/09/801,968  
; CURRENT FILING DATE: 2001-03-07  
; NUMBER OF SEQ ID NOS: 46

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: E tag
US-09-801-968-21

Query Match          25.7%; Score 70; DB 10; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.00085;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GAPVPPDPLXPR 19
Db 1 GAPVPPDPLXPR 13

RESULT 3
US-09-858-349-2
; Sequence 2, Application US/09858349
; Patent No. US20020012909A1
; GENERAL INFORMATION:
; APPLICANT: PLAKSIN, Daniel
; TITLE OF INVENTION: SMALL FUNCTIONAL UNITS OF ANTIBODY HEAVY CHAIN VARIABLE REGIONS
; FILE REFERENCE: 87534-2800
; CURRENT APPLICATION NUMBER: US/09/858,349
; CURRENT FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 136
; TYPE: PRT
; ORGANISM: mouse hybridoma specific for H-2D + RGRGRAFTVI peptide
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (99)..(107)
; OTHER INFORMATION: variable
US-09-858-349-2

Query Match          25.7%; Score 70; DB 10; Length 136;
Best Local Similarity 92.3%; Pred. No. 0.012;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GAPVPPDPLXPR 19
Db 122 GAPVPPDPLXPR 134

RESULT 4
US-09-815-242-10210
; Sequence 10210, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10210
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10210

Query Match          21.9%; Score 59.5; DB 10; Length 677;
Best Local Similarity 27.3%; Pred. No. 1.9;
Matches 12; Conservative 7; Mismatches 16; Indels 9; Gaps 1;

QY 1 YLYWIGAPVPPDPLXPRXYCXLPDCELDADHIGFOEAYRR 44
Db 250 YFYWLDAPIGY-----MGSEKNCCKRKGDSVSPDEYWK 284

RESULT 5
US-09-814-777A-15
; Sequence 15, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING TH
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU P06457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Human
US-09-814-777A-15

Query Match          20.2%; Score 55; DB 10; Length 340;
Best Local Similarity 62.5%; Pred. No. 3.6;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 YOMIGAPVPPDPLXP 18
Db 240 YGTLGTPGPPPLSP 255

RESULT 6
US-09-814-777A-18
; Sequence 18, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING TH
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU P06457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 384
; TYPE: PRT
```

```

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(482)
; NAME/KEY: NAME_feature
; LOCATION: (679)..(1919)
; OTHER INFORMATION: Exon 2
US-09-814-777A-18
```

```
Query Match          20.2%; Score 55; DB 10; Length 384;
Best Local Similarity 62.5%; Pred. No. 4.1;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Oy      3 YOMLGAPVPPDLXP 18
      | | | | | | | |
Db      284 YGTGTPGPPGPPLSP 299
```

```
RESULT 7
US-09-814-777A-20
; Sequence 20, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
```

```

; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 20
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Human
US-09-814-777A-20
```

```
Query Match          20.2%; Score 55; DB 10; Length 384;
Best Local Similarity 62.5%; Pred. No. 4.1;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Oy      3 YOMLGAPVPPDLXP 18
      | | | | | | | |
Db      284 YGTGTPGPPGPPLSP 299
```

```
RESULT 8
US-09-814-777A-100
; Sequence 100, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
```

```

; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU P06457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 100
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Human
US-09-814-777A-100
```

```
Query Match          20.2%; Score 55; DB 10; Length 470;
Best Local Similarity 62.5%; Pred. No. 5.2;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Oy      3 YOMLGAPVPPDLXP 18
      | | | | | | | |
Db      240 YGTGTPGPPGPPLSP 255
```

```
RESULT 9
US-09-815-242-11921
; Sequence 11921, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes In
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 11921
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11921
```

```
Query Match          20.0%; Score 54.5; DB 10; Length 677;
Best Local Similarity 33.3%; Pred. No. 9.1;
Matches 10; Conservative 5; Mismatches 12; Indels 3; Gaps 1;
```

```
Oy      1 YLYOMLGAPVPPDLXPRRYCXLNPPDCD 30
      | | | | | | | | : : | | |
Db      249 YFYVWLDAPRIGY--MASFRNLCARPELD 275
```

```
RESULT 10
US-09-815-242-13925
; Sequence 13925, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
```

RESULT 12  
US-09-814-777A-2  
; Sequence 2, Application US/09814777A  
; Patent No. US20020142415A1  
; GENERAL INFORMATION:  
; APPLICANT: KOOPMAN, Peter Anthony  
; APPLICANT: MUSCAT, George Eugene Orlando  
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM

```

RESULT 4
US-09-972-714-10
: Sequence 10, Application US/09972714
: Patent No. US20020106738A1
: GENERAL INFORMATION:
: APPLICANT: Fousias, George
: APPLICANT: Diamandis, Eleftherios
: TITLE OF INVENTION: NOVEL STIGLEC GENE
: FILE REFERENCE: 11757, 560501
: CURRENT APPLICATION NUMBER: US/09/972, 714
: CURRENT FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: US 60/239, 007
: PRIOR FILING DATE: 2000-10-06
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 10

```



LENGTH: 499  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-972-714-10

Query Match 19.5%; Score 53; DB 10; Length 499;  
Best Local Similarity 38.5%; Pred. No. 10;  
Matches 10; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY 2 LYQWLGAPVPPDPLXRRXYCXINP 27  
DB 188 MISWIGASVSSPGFTTARSSVLTLTP 213

RESULT 15  
US-09-765-068-2  
Sequence 2, Application US/09765068  
Patent No. US20020038009A1  
GENERAL INFORMATION:  
APPLICANT: Donoho, Gregory  
APPLICANT: Hilbun, Erin  
APPLICANT: Turner, Alex  
APPLICANT: Friedlich, Glenn  
APPLICANT: Zambrowicz, Brian  
APPLICANT: Sands, Arthur T.  
TITLE OF INVENTION: No. US20020038009A1 Human Kinase Protein and  
TITLE OF INVENTION: Polynucleotides Encoding the Same  
FILE REFERENCE: LEX-0119-USA  
CURRENT APPLICATION NUMBER: US/09/765, 068  
CURRENT FILING DATE: 2001-01-18  
PRIOR APPLICATION NUMBER: US 60/176,690  
PRIOR FILING DATE: 2000-01-18  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 422  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-765-068-2

Query Match 19.1%; Score 52; DB 10; Length 422;  
Best Local Similarity 53.8%; Pred. No. 12;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPYP 13  
DB 324 YAYDWWGRPIPTP 336

Search completed: December 4, 2002, 15:48:59  
Job time: 12 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:43:36 : Search time 141 Seconds  
(without alignments)  
224.056 Million cell updates/sec

Title: US-09-462-931-2-COPY  
Perfect score: 272  
Sequence: 1 ULYGMIGAPVPPDLXPRR.....DELADHIGFQEARFRFGPV 49

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues  
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCRTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266	97.8	49	1	PCT-US02-22821-108
2	266	97.8	49	13	US-08-973-667-1
3	266	97.8	49	13	US-08-973-667-2
4	266	97.8	49	13	US-08-973-667-3
5	266	97.8	49	13	US-08-973-667-4
6	266	97.8	49	18	US-09-462-931-2

7	266	97.8	49	25	US-10-197-954-108	Sequence 108, App
8	266	97.8	98	3	US-07-717-811A-5	Sequence 5, Appl1
9	266	97.8	98	6	US-08-246-626-5	Sequence 5, Appl1
10	266	97.8	100	1	PCT-US01-08655-186	Sequence 186, App
11	266	97.8	100	1	PCT-US01-12010-3	Sequence 3, Appl1
12	266	97.8	100	1	PCT-US01-12010-3	Sequence 23, Appl1
13	266	97.8	127	25	US-10-143-899-23	Sequence 336, App
14	266	97.8	127	26	US-10-217-651-1093	Sequence 1093, App
15	237.5	87.3	46	20	US-09-657-276-1095	Sequence 1095, App
16	222	81.6	42	14	US-09-036-085-5	Sequence 5, Appl1
17	219	80.5	140	1	PCT-US01-08655-178	Sequence 478, App
18	213	78.3	42	27	US-60-160-203-3372	Sequence 3372, App
19	213	78.3	42	27	US-60-163-123-1219	Sequence 1219, App
20	213	78.3	42	27	US-60-169-840-5089	Sequence 5089, App
21	185	68.0	43	14	US-09-036-085-5	Sequence 22, Appl
22	113	41.5	21	3	US-07-717-811A-10	Sequence 10, Appl
23	113	41.5	21	6	US-08-246-626-10	Sequence 10, Appl
24	77	28.3	15	3	US-07-717-811A-12	Sequence 12, Appl
25	77	28.3	15	6	US-08-246-626-12	Sequence 12, Appl
26	75.5	27.8	71	1	PCT-US02-30312-2166	Sequence 2166, App
27	75.5	27.8	71	1	PCT-US02-30412-2166	Sequence 2166, App
28	75.5	27.8	71	23	US-09-962-756-2166	Sequence 2166, App
29	75.5	27.8	71	26	US-10-253-471-2166	Sequence 2166, App
30	75.5	27.8	71	26	US-10-253-471-2166	Sequence 2166, App
31	75.5	27.8	73	1	PCT-US02-30312-2143	Sequence 2143, App
32	75.5	27.8	73	1	PCT-US02-30412-2143	Sequence 2143, App
33	75.5	27.8	73	23	US-09-962-756-2143	Sequence 2143, App
34	75.5	27.8	73	26	US-10-253-471-2143	Sequence 2143, App
35	75.5	27.8	73	26	US-10-253-471-2143	Sequence 2143, App
36	74.5	27.4	57	1	PCT-US02-30312-2147	Sequence 2147, App
37	74.5	27.4	57	1	PCT-US02-30412-2147	Sequence 2147, App
38	74.5	27.4	57	1	PCT-US02-30412-2147	Sequence 2147, App
39	74.5	27.4	57	23	US-09-962-756-2147	Sequence 2147, App
40	74.5	27.4	57	23	US-09-962-756-2147	Sequence 2147, App
41	74.5	27.4	57	26	US-10-253-471-2147	Sequence 2147, App
42	74.5	27.4	57	26	US-10-253-471-2147	Sequence 2147, App
43	74.5	27.4	57	26	US-10-253-471-2147	Sequence 2147, App
44	74.5	27.4	57	26	US-10-253-471-2147	Sequence 2147, App
45	74.5	27.4	57	26	US-10-253-471-2147	Sequence 2147, App

#### ALIGNMENTS

RESULT 1  
PCT-US02-22821-108  
Sequence 108, Application PC/TUS0222821  
GENERAL INFORMATION:  
APPLICANT: HK Pharmaceuticals, Inc.  
APPLICANT: Sildenafil, Subst  
APPLICANT: Little, Daniel  
TITLE OF INVENTION: Capture Compounds, Collections Thereof  
TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex  
FILE REFERENCE: 24743-2305  
CURRENT APPLICATION NUMBER: PCT/US02/22821  
CURRENT FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: 60/306,019  
PRIOR FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/314,123  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: 60/363,433  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 108  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo Sapien  
PCT-US02-22821-108

Query Match 97.8%; Score 266; DB 1; Length 49;

Best Local Similarity 93.9%; Pred. No. 3.6e-28;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 YLYOWLGAPVPPDPLXPRRYVCXLPDCELDADHIGFOEAYRRFYGPV 49  
1 YLYOWLGAPVPPDPLXPRRYVCXLPDCELDADHIGFOEAYRRFYGPV 49

RESULT 2  
US-08-973-667-1  
; Sequence 1, Application US/08973667  
; GENERAL INFORMATION:  
; APPLICANT: Sakakibara, Shunpei  
; APPLICANT: Kimura, Terutoshi  
; APPLICANT: Morimoto, Shigeto  
; TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHAYE P.C.  
; STREET: 1100 North Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/973,667  
; FILING DATE: 10-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP97/01246  
; FILING DATE: 10-APR-1997  
; APPLICATION NUMBER: JP 8-88608  
; FILING DATE: 10-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 9-43331  
; FILING DATE: 27-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 423-43  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 49 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 21  
; OTHER INFORMATION: /product= "Gla"  
US-08-973-667-1

Query Match 97.8%; Score 266; DB 13; Length 49;  
Best Local Similarity 95.9%; Pred. No. 3.6e-28;  
Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 YLYOWLGAPVPPDPLXPRRYVCXLPDCELDADHIGFOEAYRRFYGPV 49  
1 YLYOWLGAPVPPDPLXPRRYVCXLPDCELDADHIGFOEAYRRFYGPV 49

RESULT 3  
US-08-973-667-2.  
; Sequence 2, Application US/08973667  
; GENERAL INFORMATION:  
; APPLICANT: Sakakibara, Shunpei  
; APPLICANT: Kimura, Terutoshi  
; APPLICANT: Morimoto, Shigeto  
; TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHAYE P.C.  
; STREET: 1100 North Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/973,667  
; FILING DATE: 10-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP97/01246  
; FILING DATE: 10-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-88608  
; FILING DATE: 10-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 9-43331  
; FILING DATE: 27-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 423-43  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 49 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 17  
; OTHER INFORMATION: /product= "Gla"  
US-08-973-667-2

Query Match 97.8%; Score 266; DB 13; Length 49;  
Best Local Similarity 98.0%; Pred. No. 3.6e-28;  
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YLYOWLGAPVPPDPLXPRRYVCXLPDCELDADHIGFOEAYRRFYGPV 49  
1 YLYOWLGAPVPPDPLXPRRYVCXLPDCELDADHIGFOEAYRRFYGPV 49

RESULT 4  
US-08-973-667-3  
; Sequence 3, Application US/08973667  
; GENERAL INFORMATION:

```

: APPLICANT: Sakakibara, Shunpei
: APPLICANT: Kimura, Terutoshi
: APPLICANT: Morimoto, Shigeto
: TITLE OF INVENTION: ANTI-GLI017-OSTEOCALCIN ANTIBODY
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHAYE P.C.
: STREET: 1100 North Glebe Rd. 8th floor
: CITY: Arlington
: STATE: VA
: COUNTRY: USA
: ZIP: 22201-4741
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/973,667
: FILING DATE: 10-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP97/01246
: FILING DATE: 10-APR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-88608
: FILING DATE: 10-APR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 9-43331
: FILING DATE: 27-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.
: REGISTRATION NUMBER: 25,327
: REFERENCE/DOCKET NUMBER: 423-43
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-816-4000
: TELEFAX: 703-816-4100
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Human
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 17
: OTHER INFORMATION: /product= "Gla"
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 21
: OTHER INFORMATION: /product= "Gla"
: US-08-973-667-3
:
: Query Match 97.8%; Score 266; DB 13; Length 49;
: Best Local Similarity 98.0%; Pred. No. 3.6e-28;
: Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
:
: Oy 1 YLYWMLGAPVYPDPLXPRXVCXLPNDCDELADHIGFOEAYRRFGPV 49
: 1 YLYWMLGAPVYPDPLXPRXVCXLPNDCDELADHIGFOEAYRRFGPV 49
:
: RESULT 5
: US-08-973-667-4
: Sequence 4, Application US/08973667
: GENERAL INFORMATION:
: APPLICANT: Sakakibara, Shunpei
: APPLICANT: Kimura, Terutoshi
: APPLICANT: Morimoto, Shigeto
: TITLE OF INVENTION: ANTI-GLI017-OSTEOCALCIN ANTIBODY

```

```

: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHAYE P.C.
: STREET: 1100 North Glebe Rd. 8th floor
: CITY: Arlington
: STATE: VA
: COUNTRY: USA
: ZIP: 22201-4741
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/973,667
: FILING DATE: 10-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP97/01246
: FILING DATE: 10-APR-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-88608
: FILING DATE: 10-APR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 9-43331
: FILING DATE: 27-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.
: REGISTRATION NUMBER: 25,327
: REFERENCE/DOCKET NUMBER: 423-43
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-816-4000
: TELEFAX: 703-816-4100
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Human
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 17
: OTHER INFORMATION: /product= "Gla"
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 21
: OTHER INFORMATION: /product= "Gla"
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 24
: OTHER INFORMATION: /product= "Gla"
: US-08-973-667-4
:
: Query Match 97.8%; Score 266; DB 13; Length 49;
: Best Local Similarity 100.0%; Pred. No. 3.6e-28;
: Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: Oy 1 YLYWMLGAPVYPDPLXPRXVCXLPNDCDELADHIGFOEAYRRFGPV 49
: 1 YLYWMLGAPVYPDPLXPRXVCXLPNDCDELADHIGFOEAYRRFGPV 49
:
: RESULT 6
: US-09-462-931-2
: Sequence 2, Application US/09462931
: GENERAL INFORMATION:
: APPLICANT: HELLMAN, Jukka
: APPLICANT: KRTNEN, Sanna-Maria
: APPLICANT: KARP, Matti
: APPLICANT: LTVGREN, Timo

```

```

? APPLICANT: VNNEN, Kalervo
? APPLICANT: PETERSSON, Kim
? TITLE OF INVENTION: Isolated osteocalcin fragments
? FILE REFERENCE: Isolated osteocalcin fragments
? CURRENT APPLICATION NUMBER: US/09/462,931
? CURRENT FILING DATE: 2000-01-18
? PRIOR APPLICATION NUMBER: PCT/FI98/00550
? PRIOR FILING DATE: 1998-06-24
? PRIOR APPLICATION NUMBER: FI 973371
? PRIOR FILING DATE: 1997-08-15
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: patentln Ver. 2.1
? SEQ ID NO 2
? LENGTH: 49
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: peptide
? LOCATION: (1)..(49)
? OTHER INFORMATION: Glu at residues 17, 21 and 24 may be gamma-carboxy-Glu
US-09-462-931-2

```

Query Match 97.8%; Score 266; DB 3; Length 98;  
Best Local Similarity 93.9%; Pred. No. 7.8e-28;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YLYQWLGAPVPPDPLXPRXVCXLNPDCELADHIGFOEAYRRFGPV 49  
|||||  
Db 50 YLYQWLGAPVPPDPLXPRXVCXLNPDCELADHIGFOEAYRRFGPV 98

## RESULT 9

US-08-246-626-5  
; Sequence 5, Application US/08246626  
; GENERAL INFORMATION:  
; APPLICANT: Hiroshi EGUCHI et al.  
; TITLE OF INVENTION: Recombinant Human Osteocalcin  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; City: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: diskette, 5.25 inch, 500 Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/246,626  
; FILING DATE: 20-MAY-1994  
; CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/717,811  
; FILING DATE: 19-Jun-1991  
; APPLICATION NUMBER: US 08/131,932  
; FILING DATE: 05-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 98 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:

## FEATURE:

;; NAME/KEY:  
;; LOCATION:  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION:  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; PAGES:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO:  
US-08-246-626-5

Query Match 97.8%; Score 266; DB 6; Length 98;  
Best Local Similarity 93.9%; Pred. No. 7.8e-28;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YLYQWLGAPVPPDPLXPRXVCXLNPDCELADHIGFOEAYRRFGPV 49  
|||||  
Db 50 YLYQWLGAPVPPDPLXPRXVCXLNPDCELADHIGFOEAYRRFGPV 98

## RESULT 10

PCT-US01-08655-186  
; Sequence 186, Application PC/TUS0108655  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-065  
; CURRENT APPLICATION NUMBER: PCT/US01/08655  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 09/522,929  
; PRIOR FILING DATE: 2000-04-18  
; PRIOR APPLICATION NUMBER: 09/770,160  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/668,317  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 09/695,783  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: 09/728,628  
; PRIOR FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: 09/783,066  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 09/816,828  
; PRIOR FILING DATE: 2001-03-22  
; NUMBER OF SEQ ID NOS: 584  
; SOFTWARE: Custom  
; SEQ ID NO 186  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-08655-186

Query Match 97.8%; Score 266; DB 1; Length 100;  
Best Local Similarity 93.9%; Pred. No. 8e-28;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YLYQWLGAPVPPDPLXPRXVCXLNPDCELADHIGFOEAYRRFGPV 49  
|||||  
Db 52 YLYQWLGAPVPPDPLXPRXVCXLNPDCELADHIGFOEAYRRFGPV 100

## RESULT 11

PCT-US01-12010-3  
; Sequence 3, Application PC/TUS0112010  
; GENERAL INFORMATION:  
; APPLICANT: Genaisance Pharmaceuticals, Inc.

APPLICANT: Bentivegna, Steven C.  
APPLICANT: Chew, Anne  
APPLICANT: Choi, Julie Y.  
APPLICANT: Koshy, Beena  
APPLICANT: Rounds, Eileen  
APPLICANT: Stephens, J. Claiborne  
TITLE OF INVENTION: Haplotypes of the BGLAP Gene  
FILE REFERENCE: MMH-0512PCT BGLAP  
CURRENT APPLICATION NUMBER: PCT/US01/12010  
CURRENT FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 60/195,840  
PRIOR FILING DATE: 2000-04-11  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Homo sapien  
PCT-US01-12010-3

Query Match 97.8%; Score 266; DB 1; Length 100;  
Best Local Similarity 93.9%; Pred. No. 8e-28; Indels 0; Gaps 0;  
Matches 46; Conservative 0; Mismatches 3;

Qy 1 YLYQWLGAVPYPDPLPRXYCXLPDCDELADHIGFOEAYRREYGPV 49  
|||||  
Db 52 YLYQWLGAVPYPDPLPRXYCXLPDCDELADHIGFOEAYRREYGPV 100

RESULT 12  
US-10-143-899-23  
Sequence 23, Application US/10143899  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT247CIN  
CURRENT APPLICATION NUMBER: US/10/143,899  
CURRENT FILING DATE: 2002-05-14  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 23  
LENGTH: 127  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (18)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (64)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-143-899-23

Query Match 97.8%; Score 266; DB 25; Length 127;  
Best Local Similarity 93.9%; Pred. No. 1e-27; Indels 0; Gaps 0;  
Matches 46; Conservative 0; Mismatches 3;

Qy 1 YLYQWLGAVPYPDPLPRXYCXLPDCDELADHIGFOEAYRREYGPV 49  
|||||  
Db 79 YLYQWLGAVPYPDPLPRXYCXLPDCDELADHIGFOEAYRREYGPV 127

RESULT 13  
US-10-217-651-336  
Sequence 336, Application US/10217651  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: P1211CIN  
CURRENT APPLICATION NUMBER: US/10/217,651  
CURRENT FILING DATE: 2002-08-14

PRIOR APPLICATION NUMBER: 09/760,491  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/214,886  
PRIOR FILING DATE: 2000-06-28  
PRIOR APPLICATION NUMBER: 60/217,487  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,758  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,963  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/217,496  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,447  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/218,290  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/225,757  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/226,868  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/216,647  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/225,267  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/216,880  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/225,270  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/251,869  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/235,834  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/234,274  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/234,223  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/228,924  
PRIOR FILING DATE: 2000-08-30  
PRIOR APPLICATION NUMBER: 60/224,518  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/236,369  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/224,519  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,964  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/241,809  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/249,299  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/236,327  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/241,785  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/244,617  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 60/225,268  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/236,368  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/251,856  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/251,868  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/229,344  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/234,997  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: 60/229,343



PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229, 345  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229, 287  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229, 513  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/231, 413  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/229, 509  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/236, 367  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/237, 039  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237, 038  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/236, 370  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/236, 802  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237, 037  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237, 040  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/240, 960  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/239, 935  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/239, 937  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/241, 787  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/246, 474  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/246, 532  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/249, 216  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 210  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/226, 681  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/225, 759  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/225, 213  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/227, 182  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/225, 214  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/235, 836  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/230, 438  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/215, 135  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: 60/225, 266  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/249, 218  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 208  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 213  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 212  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 207  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 245  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 244  
PRIOR FILING DATE: 2000-11-17

PRIOR APPLICATION NUMBER: 60/249, 217  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 211  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 215  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 264  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 214  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249, 297  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/232, 400  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/231, 242  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/232, 081  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/232, 080  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/231, 414  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/231, 244  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/233, 064  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/233, 063  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232, 397  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232, 399  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232, 401  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/241, 808  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241, 826  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241, 786  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241, 221  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/246, 475  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/231, 243  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/233, 065

Query Match 97.8%; Score 266; DB 26; Length 127;  
Best Local Similarity 93.9%; Pred. No. 1e-27; Indels 0; Gaps 0;  
Matches 46; Conservative 0; Mismatches 3;

OY 1 YLYQWLGAPVPPDPLXPRRYCXLNPDDELADHIGFOEAYRRFGPV 49  
DB 79 YLYQWLGAPVPPDPLXPRRYCXLNPDDELADHIGFOEAYRRFGPV 127

RESULT 14  
US-09-657-276-1093  
Sequence 1093, Application US/09657276  
GENERAL INFORMATION:  
APPLICANT: Conjuchem, Inc.  
APPLICANT: Bridon, Dominique  
APPLICANT: Ezrin, Alan  
APPLICANT: Milner, Peter  
APPLICANT: Holmes, Darren  
APPLICANT: Thibaudau, Karen  
TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM  
PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD  
FILE REFERENCE: 2110  
CURRENT APPLICATION NUMBER: US/09/657, 276  
CURRENT FILING DATE: 2000-09-07

; PRIOR APPLICATION NUMBER: 60/134,406  
; PRIOR FILING DATE: 1999-05-17  
; PRIOR APPLICATION NUMBER: 60/153,406  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: 60/159,783  
; PRIOR FILING DATE: 1999-10-18  
; NUMBER OF SEQ ID NOS: 1617  
; SOFTWARE: Patentl Ver. 2.1  
; SEQ ID NO 1093  
; LENGTH: 47  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-657-276-1093

Query Match 90.8%; Score 247; DB 20; Length 47;  
Best Local Similarity 93.9%; Pred. No. 1.3e-25;  
Matches 46; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Oy 1 YLYQWLGAIPVPPDLPXRRXYCXLNPDDELADHIGFOEAYRRFGPV 49  
|||||  
Db 1 YLYQWLGAIPVPPDLPXRRXYCXLNPDDELADHIGFOEAYRRFGPV 47

RESULT 15  
US-09-657-276-1095  
; Sequence 1095, Application US/09657276  
; GENERAL INFORMATION:  
; APPLICANT: Conjuchem, Inc.  
; APPLICANT: Bridon, Dominique  
; APPLICANT: Ezrin, Alan  
; APPLICANT: Milner, Peter  
; APPLICANT: Holmes, Darren  
; APPLICANT: Thibaudau, Karen  
; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM  
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD  
; FILE REFERENCE: 2110  
; CURRENT APPLICATION NUMBER: US/09/657,276  
; PRIOR FILING DATE: 2000-09-07  
; PRIOR APPLICATION NUMBER: 60/134,406  
; PRIOR FILING DATE: 1999-05-17  
; PRIOR APPLICATION NUMBER: 60/153,406  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: 60/159,783  
; PRIOR FILING DATE: 1999-10-18  
; NUMBER OF SEQ ID NOS: 1617  
; SOFTWARE: Patentl Ver. 2.1  
; SEQ ID NO 1095  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-657-276-1095

Query Match 87.3%; Score 237.5; DB 20; Length 46;  
Best Local Similarity 93.9%; Pred. No. 2.6e-24;  
Matches 46; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Oy 1 YLYQWLGAIPVPPDLPXRRXYCXLNPDDELADHIGFOEAYRRFGPV 49  
|||||  
Db 1 YLYQWLGAIPVPPDLPXRRXYCXLNPDDELADHIGFOEAYRRFGPV 46

Search completed: December 4, 2002, 15:48:01  
Job time : 142 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:44:36 : Search time 13 Seconds  
(Without alignments)  
230.896 Million cell updates/sec

Title: US-09-462-931-2-COPY

Sequence: 1 YLYQWLGAAPVPPDLXPRR.....DELADHIGFOEAYRRFGPV 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193982 seqs, 61258239 residues

Total number of hits satisfying chosen parameters: 193982

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266	97.8	100	6 US-10-283-656-1	Sequence 1, Appl 1
2	224	82.4	58	5 US-09-724-676-90318	Sequence 90318, A
3	224	82.4	58	5 US-09-724-676A-90318	Sequence 90318, A
4	70	25.7	13	5 US-09-802-154-21	Sequence 21, Appl 1
5	70	25.7	257	6 US-10-096-246-2	Sequence 2, Appl 1
6	58	21.3	92	1 PCT-US02-32727-11781	Sequence 11781, A
7	57.5	21.1	194	5 US-09-724-676-93282	Sequence 93282, A
8	57.5	21.1	194	5 US-09-724-676-93283	Sequence 93283, A
9	57.5	21.1	194	5 US-09-724-676-93284	Sequence 93284, A
10	57.5	21.1	194	5 US-09-724-676-93285	Sequence 93285, A
11	57.5	21.1	194	5 US-09-724-676-93286	Sequence 93286, A
12	57.5	21.1	194	5 US-09-724-676-93287	Sequence 93287, A
13	57.5	21.1	194	5 US-09-724-676-93288	Sequence 93288, A
14	57.5	21.1	194	5 US-09-724-676-93290	Sequence 93290, A
15	57.5	21.1	194	5 US-09-724-676-93291	Sequence 93291, A
16	57.5	21.1	194	5 US-09-724-676-93292	Sequence 93292, A
17	57.5	21.1	194	5 US-09-724-676-93293	Sequence 93293, A
18	57.5	21.1	194	5 US-09-724-676-93294	Sequence 93294, A
19	57.5	21.1	194	5 US-09-724-676-93295	Sequence 93295, A
20	57.5	21.1	194	5 US-09-724-676-93296	Sequence 93296, A
21	57.5	21.1	194	5 US-09-724-676-93297	Sequence 93297, A
22	57.5	21.1	194	5 US-09-724-676-93298	Sequence 93298, A
23	57.5	21.1	194	5 US-09-724-676-93299	Sequence 93299, A
24	57.5	21.1	194	5 US-09-724-676-93301	Sequence 93301, A
25	57.5	21.1	194	5 US-09-724-676A-93282	Sequence 93282, A
26	57.5	21.1	194	5 US-09-724-676A-93283	Sequence 93283, A

27	57.5	21.1	194	5 US-09-724-676A-93284	Sequence 93284, A
28	57.5	21.1	194	5 US-09-724-676A-93285	Sequence 93285, A
29	57.5	21.1	194	5 US-09-724-676A-93286	Sequence 93286, A
30	57.5	21.1	194	5 US-09-724-676A-93287	Sequence 93287, A
31	57.5	21.1	194	5 US-09-724-676A-93288	Sequence 93288, A
32	57.5	21.1	194	5 US-09-724-676A-93290	Sequence 93290, A
33	57.5	21.1	194	5 US-09-724-676A-93291	Sequence 93291, A
34	57.5	21.1	194	5 US-09-724-676A-93292	Sequence 93292, A
35	57.5	21.1	194	5 US-09-724-676A-93293	Sequence 93293, A
36	57.5	21.1	194	5 US-09-724-676A-93294	Sequence 93294, A
37	57.5	21.1	194	5 US-09-724-676A-93295	Sequence 93295, A
38	57.5	21.1	194	5 US-09-724-676A-93296	Sequence 93296, A
39	57.5	21.1	194	5 US-09-724-676A-93297	Sequence 93297, A
40	57.5	21.1	194	5 US-09-724-676A-93298	Sequence 93298, A
41	57.5	21.1	194	5 US-09-724-676A-93299	Sequence 93299, A
42	57.5	21.1	194	5 US-09-724-676A-93301	Sequence 93301, A
43	57.5	21.1	253	5 US-09-724-676-93242	Sequence 93242, A
44	57.5	21.1	253	5 US-09-724-676-93243	Sequence 93243, A
45	57.5	21.1	253	5 US-09-724-676-93244	Sequence 93244, A

## ALIGNMENTS

RESULT 1  
US-10-283-656-1  
: Sequence 1, Application US/10283656  
: GENERAL INFORMATION:  
: APPLICANT: EKEMA, George Mbella  
: APPLICANT: MAVS, Robert W.  
: APPLICANT: BRUNDEN, Kurt R.  
: TITLE OF INVENTION: Methods for Using Osteocalcin  
: FILE REFERENCE: ATX-005  
: CURRENT APPLICATION NUMBER: US/10/283,656  
: CURRENT FILING DATE: 2002-10-29  
: NUMBER OF SEQ ID NOS: 4  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 1  
: LENGTH: 100  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
: US-10-283-656-1

Query Match 97.8%; Score 266; DB 6; Length 100;  
Best Local Similarity 93.9%; Pred. No. 1.2e-27;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 YLYQWLGAAPVPPDLXPRRXVXILNPDDELADHIGFOEAYRRFGPV 49  
Db 52 YLYQWLGAAPVPPDLXPRRXVXILNPDDELADHIGFOEAYRRFGPV 100

RESULT 2  
US-09-724-676-90318  
: Sequence 90318, Application US/09724676  
: GENERAL INFORMATION:  
: APPLICANT: CompuGen LTD  
: TITLE OF INVENTION: Variants of alternative splicing  
: FILE REFERENCE: 129181.4 CompuGen  
: CURRENT APPLICATION NUMBER: US/09/724,676  
: CURRENT FILING DATE: 2000-11-28  
: NUMBER OF SEQ ID NOS: 97222  
: SOFTWARE: PatentIn version 3.2  
: SEQ ID NO 90318  
: LENGTH: 58  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
: US-09-724-676-90318

Query Match 82.4%; Score 224; DB 5; Length 58;  
Best Local Similarity 90.9%; Pred. No. 2e-22;  
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY      6 LGAPVPPDPLAPRRVXCXLPDDELADHIGFQEA YRRFYGPV 43
      | | | | | | | | | | | | | | | | | | | | | |
Db     15 LRAVPYPDPLEPRREVCEINPDDELADHIGFQEA YRRFYGPV 58

```

RESULT 3  
US-09-724-676A-90318  
; Sequence 90318, AP

```

: Sequence 90318, Application US/09724676A
:
: GENERAL INFORMATION:
:
: APPLICANT: CompuGen LTD
: TITLE OF INVENTION: Variants of alternative splicing
: FILE REFERENCE: 129181.4 CompuGen
: CURRENT APPLICATION NUMBER: US/09/724,676A
: CURRENT FILING DATE: 2000-11-28
: NUMBER OF SEQ ID NOS: 97222
:
: SOFTWARE: PatentIn version 3.2

```

ORGANISM: Homo sapiens  
US-09-724-676A-90318

Query Match	82.48;	Score 224;	DB 5;	Length 58;
Best Local Similarity	90.98;	Pred. No. 2e-22;		
Matches	40;	Conservative	0;	Mismatches 4;
				Indels

OY      6 LGAPVYPDPLPRRXYCLNPDCDELADHIGFQEAIRRFYGPV 49  
          | | | | | | | | | | | | | | | | | | | | | |  
DB     15 LRAVPYPDPLEPREVCNELNPDCELADHIGFQEAIRRFYGPV 58

RESULT 4  
US-09-802-154-21  
; Sequence 21, Application US/09802154

```

1  APPLICANT: Itoh, Nobuyuki, Michael
2  APPLICANT: Kavanaugh, W. Michael
3  TITLE OF INVENTION: HUMAN RCF-23 GENE AND GENE EXPRESSION
4  TITLE OF INVENTION: PRODUCTS
5  FILE REFERENCE: PP-17149, .001/201130.409
6  CURRENT APPLICATION NUMBER: US/09/802,154
7  CURRENT FILING DATE: 2001-03-07
8  NUMBER OF SEQ ID NOS: 46
9  SOFTWARE: FastSeq for Windows Version 4.0
10 SEQ ID NO 21
11 LENGTH: 13
12 TYPE: PRT
13 ORGANISM: Artificial Sequence
14 FEATURE:
15 OTHER INFORMATION: E tag
16
17 US-09-802-154-21

```

Query Match	25.7%	Score 70:	DB 5:	Length 13:
Best Local Similarity	92.3%	Pred. No.	0.0056:	
Matches 12; Conservative	0:	Mismatches 1:	Indels 0:	Gaps 0:

QY	7	GAPVPYPDP LKPR	19
Db	1	GAPVPYPDP LKPR	13

RESULT 5  
US-10-096-246-2  
; Sequence 2, A

```
; Sequence 2, Application US/10096246
; GENERAL INFORMATION:
; APPLICANT: The Minister of National Defence, Government of Canada
```

```

; APPLICANT: Alvi, Azhar E
; APPLICANT: Nagata, Leslie

```

TITLE OF INVENTION:	Clothing, Expression, Sequencing, and Functional Enhancement of M
TITLE OF INVENTION:	scFv Antibody Against Venezuelan Equine Encephalitis Virus (Vee)

```

: CURRENT FILING DATE: 2002-03-13
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 2
: LENGTH: 257
: TYPE: PRT
: ORGANISM: Mouse hybridoma cell line 1A4A1
: US-10-036-246-2

```

Query Match	25.7%	Score 70;	DB 6;	length 257;
Best Local Similarity	92.3%	Pred. No. 0.11;		
Matches 12; Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;

```
QY      7 GAPVPYPDDLXPR 19
          |||||
Db     243 GAPVPYPDDLXPR 255
```

RESULT 6  
PCT-US02-32727-11781

```
; Sequence 11781, Application PC/TUS0232727
; GENERAL INFORMATION:
```

APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir

APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay

```

; APPLICANT:  Maisonneuve, Jean Francois
; APPLICANT:  Zhang, Yanni

```

APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shyian

APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT:

APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darrick  
; APPLICANT:

APPLICANT: Barth, Brenda  
; APPLICANT: Douglass, John  
; APPLICANT: Douglass, John

;; TITLE OF INVENTION: Compositions and Methods  
;; FILE REFERENCE: 210121.514C1  
;;

; CURRENT APPLICATION NUMBER: PCT/US02/32727  
 ; CURRENT FILING DATE: 2002-10-11  
 ; CURRENT PCT NO.: 2002

```

; NUMBER OF SEQ
; SEQ ID NO 11781
;

```

```

; LENGTH: 92
; TYPE: PRT
;

```

ORGANISM: Propioni acnes  
PCT-US02-32727-11781

Query Match	21.3%;	Score 58;	DB 1;	Length 92;
Best Local Similarity	31.2%;	Pred. No. 1.5;		
Matches 15;	Conservative	5;	Mismatches 28;	Indels 0;
				Gaps 0;

```
OY      1 YLYQWLGAFPVPPDELPXRRXVCLNPDCELDADHIGFQEAVERREYGP 48
      :| | | : | | | : |
Db      15 FLVWMLVAPILRLSLPVTPRRLVSMASPTPSKCRITWCSSMAAVRPVYPP 62
```

RESULT 7  
US-09-724-676-93282

```
; GENERAL INFORMATION:  
; APPLICANT: CompuGen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
FILED NOVEMBER 19 1987 AT COMPUGEN
```

```

; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28

```

```

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
;

```

```

; SOFTWARE: Pate
; SEQ ID NO 93282
;

```

```

; LENGTH: 194
; TYPE: PRT
; OBJECTIVE: homo sapiens

```

ORGANISM: HOMO  
US-09-724-676-93282





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:43:01 : Search time 19 seconds  
(without alignments)  
247,926 Million cell updates/sec

Title: US-09-462-931-2-COPY  
Perfect score: 272  
Sequence: 1 YLYQWLGAVPYPPDPLXPRR.....DELADHIGQEAAYRRYGPV 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	266	97.8	100	1	GEHU osteocalcin precur
2	239	95.2	49	1	GEHU osteocalcin - crab
3	244	89.7	100	1	GEBO osteocalcin precur
4	219	80.5	49	1	GECT osteocalcin - cat
5	202	74.3	99	1	GEPT osteocalcin precur
6	199	73.2	49	1	AG1280 osteocalcin - rabb
7	181	66.5	48	1	S02208 osteocalcin - emu
8	168.5	61.9	97	1	GECH osteocalcin precur
9	148	54.4	95	2	B25471 osteocalcin precur
10	148	54.4	95	2	I53275 osteocalcin - mus
11	144	52.9	95	2	I67413 osteocalcin - mus
12	144	52.9	95	2	I61188 osteocalcin - relate
13	91	33.5	47	1	GEHF osteocalcin - swor
14	68	25.0	45	1	A42794 osteocalcin - blue
15	59.5	21.9	677	1	SYECMT methionine-trna 11
16	59.5	21.9	677	2	H90993 methionine trna sy
17	59.5	21.9	677	2	C85839 methionine trna sy
18	59	21.7	706	2	T15701 hypothetical prote
19	57.5	21.1	338	2	G01496 transcripition fact
20	57.5	21.1	363	2	I38937 DNA/RNA-binding pr
21	57.5	21.1	430	2	D64373 hypothetical prote
22	57	21.0	574	2	B88465 protein B0244.8 [1
23	56.5	20.8	373	2	S32537 erythroid transcri
24	56.5	20.8	435	2	S52784 ornithine decarbox
25	56	20.6	262	2	S45026 ribosomal protein
26	56	20.6	264	2	S47642 ribosomal protein
27	56	20.6	536	2	T42606 probable transcrip
28	55	20.2	271	2	S76939 hypothetical prote
29	55	20.2	516	2	T15633 hypothetical prote

30	54.5	20.0	139	2	S65969	YYCE protein - Bac
31	54.5	20.0	677	2	C83210	methionyl-trna syn
32	54.5	20.0	677	2	AH0776	methionine-trna 11
33	54.5	20.0	699	2	AG0371	probable acetyltra
34	54	19.9	227	2	T06624	hypothetical prote
35	54	19.9	265	2	T01187	ribosomal protein
36	54	19.9	527	2	H85135	hypothetical prote
37	53.5	19.7	486	1	VCLJH2	env polyprotein -
38	53.5	19.7	791	2	G81109	ribonuclease II fa
39	53	19.5	285	2	T01203	ribosomal protein
40	53	19.5	291	2	A43674	US2 protein - huma
41	53	19.5	415	2	S55617	hypothetical prote
42	53	19.5	726	1	S73915	virulence-associat
43	53	19.5	1344	2	T14313	rig-1 protein - mo
44	53	19.5	2257	2	D86483	protein F535.19 [1
45	52.5	19.3	221	1	C69009	conserved hypothet

#### ALIGNMENTS

##### RESULT 1

GEHU

osteocalcin precursor (validated) - human

N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing prote

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1980 #sequence-revision 07-Apr-1994 #text-change 08-Dec-2000

C:Accession: S12652; C25471; A03301; S08694

R:Kiefer, M.C.; Saphire, A.C.S.; Bauer, D.M.; Barr, P.J.

Nucleic Acids Res. 18, 1909, 1990

A>Title: The cDNA and derived amino acid sequences of human and bovine bone gla prote

A:Reference number: S12652; MID:90245603; PMID:2336375

A:Accession: S12652

A:Molecule type: mRNA

A:Residues: 1-100 <KIE>

A:Cross-references: EMBL:X53698; NID:936092; PIDN:CAA37736.1; PID:936093

R:Celeste, A.J.; Rosen, V.; Buckner, J.L.; Kitz, R.; Wang, E.A.; Wozney, J.M.

EMO J. 5, 1985-1890, 1986

A>Title: Isolation of the human gene for bone gla protein utilizing mouse and rat cDN

A:Reference number: A91045; MID:87004555; PMID:3019668

A:Accession: C25471

A:Molecule type: DNA

A:Residues: 1-32, 35-100 <CELE>

A:Cross-references: EMBL:X04143; NID:929449; PIDN:CAA27763.1; PID:929450

R:Poser, J.W.; Esch, F.S.; Ling, N.C.; Price, P.A.

J. Biol. Chem. 255, 8685-8691, 1980

A>Title: Isolation and sequence of the vitamin K-dependent protein from human bone.

A:Reference number: A03301; MID:81006914; PMID:6967872

A:Accession: A03301

A:Molecule type: protein

A:Residues: 52-100 <POS>

R:Calins, J.R.; Williamson, M.K.; Price, P.A.

Anal. Biochem. 199, 93-97, 1991

A>Title: Direct identification of gamma-carboxyglutamic acid in the sequencing of vit

A:Reference number: A44566; MID:92222128; PMID:1807167

A:Contents: annotation

C:Comment: This protein, isolated from bone, binds strongly to apatite.

C:Comment: Alternative splicing may produce the sequence presented in reference A9104

C:Comment: Glu-68 is gamma-carboxylated in 9-50% of the molecules.

C:Genetics:

A:Gene: GDB:BGAP

A:Cross-references: GDB:118760; OMIM:112260

A:Map position: 1925-1931

A:Introns: 22/1; 35/1; 58/2

C:Superfamily: osteocalcin

C:Keywords: bone; calcium binding; carboxyglutamic acid; extracellular matrix

F:1-51/Domain: signal sequence #status predicted <SIG>

F:52-100/Product: osteocalcin #status experimental <MAT>

F:50/Modified site: 4-hydroxyproline (Pro) #status absent

F:58/Modified site: gamma-carboxyglutamic acid (Glu) (partial) #status experimental

F:72/75/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental

F:74-80/Disulfide bonds: #status experimental

Query Match	97.8%	Score 266	DB 1	Length 100
Best Local Similarity	93.9%	Pred. No. 1,1e-27		
Matches 46	Conservative	0	Mismatches 3	Indels 0
			Gaps 0	
QY	1	YLQWLGAPVPPDPLPRRYXCXNPPCDELADHIGQEAATRRYGPV	49	
Db	52	YLQWLGAPVPPDPLPRRYXCXNPPCDELADHIGQEAATRRYGPV	100	

## RESULT 2

osteocalcin - crab-eating macaque  
N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing protein  
C:Species: *Macaca fascicularis* (crab-eating macaque)  
C:Date: 27-Nov-1965 #sequence\_revision 27-Nov-1965 #text\_change 06-Sep-1996  
C:Accession: A03302  
R:Häuschka, P.V.; Carr, S.A.; Biemann, K.  
Biochemistry 21, 638-642, 1982  
A:Title: Primary structure of monkey osteocalcin.  
A:Reference number: A03302; PMID:82182842; PMID:6978733  
A:Accession: A03302  
A:Molecule type: protein  
A:Residues: 1-49 <HAU>  
C:Comment: This protein, isolated from bone, binds strongly to apatite.  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
F:9/Modified site: 4-hydroxyproline (Pro) #status experimental  
F:17/21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:23-29/Disulfide bonds: #status experimental

Query Match

Query Match	95.2%	Score 259;	DB 1;	Length 49;
Best Local Similarity	89.8%	Pred. NO. 4.1e-27;		
Matches 44;	Conservative 1;	Mismatches 4;	Indels 0;	Gaps 0;

**QY**

1 ILYQWLGA~~PVPPDPLXPRRYXC~~LNPDCDELADHIGFQEAIRRRYPV 493  
||||||| : |||||  
1 ILYQWLGAPARPPDPLEPKREVCELNPDCEDELADHIGFQEAIRRRYPV 499

**Dd**

### RESULT 3

osteocalcin precursor - bovine  
N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing protein  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 24-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 22-Jun-1999  
C:Accession: S12653; A03303; S08693  
R:Kiefer, M.C.; Saphire, A.C.; Bauer, D.M.; Barr, P.J.  
Nucleic Acids Res. 18, 1909, 1990  
A:Title: The cDNA and derived amino acid sequences of human and bovine bone Gla protein  
A:Reference number: S12652; MUID:90245603; PMID:2336375  
A:Accession: S12653  
A:Molecule type: mRNA  
A:Residues: 1-100 <KIE>  
A:Cross-references: EMBL:X53699; NID:G719; PIDN:CA37737.1; PID:G720  
A:Notes: alternative splicing may produce a sequence lacking residues 33-34  
R:Price, P.A.; Poser, J.W.; Raman, N  
Proc. Natl. Acad. Sci. U.S.A. 73, 3374-3375, 1976  
A:Title: Primary structure of the gamma-carboxyglutamic acid-containing protein from bovine bone  
A:Reference number: A03303; MUID:77036749; PMID:1068450  
A:Accession: A03303  
A:Molecule type: Protein  
A:Residues: 52-100 <PRI>  
C:Comment: This protein, isolated from bone, binds strongly to apatite.  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
F:1-51/Domain: signal sequence #status predicted <SIG>  
F:52-100/Product: osteocalcin #status experimental <MAT>  
F:60/Modified site: 4-hydroxyproline (Pro) #status experimental  
F:68,72,75/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:74-80/Disulfide bonds: #status experimental

Query Match	89.7%	Score 244;	DB 1;	Length 100;
Best Local Similarity	85.7%	Pred. No. 8.1e-25;		

Best Local Similarity 85.7%; Pred. No. 8.1e-25;

Matches 42; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPYDPLXPRRXVCXLNPDCELDADHIGFQEAAYRRFYGPV 49  
|| ||||| : |||||  
Db 52 YLDHWLGAAPAPYDPLPKREVCELNPDCELDADHIGFQEAAYRRFYGPV 100

## RESULT 4

osteocalcin - cat  
 N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: *Felis silvestris catus* (domestic cat)  
 C:Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 06-Sep-1996  
 C:Accession: A03304  
 R:Shimomura, H.; Kanai, Y.; Sanada, K.  
 J. Biochem. 96, 405-411, 1984  
 A:Title: Primary structure of cat osteocalcin.  
 A:Reference number: A03304; MUID:85054706; PMID:6334077  
 A:Accession: A03304  
 A:Molecule type: protein  
 A:Residues: 1-49 <SH3>  
 C:Superfamily: osteocalcin  
 C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
 F:9,Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:17,21,24,Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:23-29,Disulfide bonds: #status experimental

Query Match

Query Match	80.5%	Score 219	DB 1	Length 49
Best Local Similarity	77.6%	Pred. No. 7.2e-22		
Matches 38; Conservative	3;	Mismatches 8;	Indels 0;	Gaps 0

```
QY      1 ILYWLGAPVPPDPLXPRXYCXLPDCDELADHIGFQEAYRRRYGPV 499
        || ||| |||| : : ||||| ||||| : ||||| |
Db      1 YLAPGLGAPAPYPDPLEPKREICELNPDCELADHIGFGDAYRRRGTV 499
```

## RESULT 5

Osteocalcin precursor - rat  
 A:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 30-Sep-1987 #sequence,revision 30-Sep-1987 #text,change 22-Jun-1999  
 C:Accession: A31856; A31419; A32324; A25167; A25471  
 R:Yoon, K.; Rutledge, S.J.C.; Buenaga, R.F.; Rodan, G.A.  
 Biochemistry 27, 8521-8526, 1988  
 A:Title: Characterization of the rat osteocalcin gene: stimulation of promoter activity  
 A:Reference number: A31856; MUID:89118266; PMID:3265336  
 A:Accession: A31856  
 A:Molecule type: DNA  
 A:Residues: 1-99 <TOD>  
 A:Cross-references: GB:M23637; NID:g340986; PIDN:AAA41761.1; PID:g514962  
 R:Theofan, G.; Haberstroh, L.M.; Price, P.A.  
 DNA 8, 213-221, 1989  
 A:Title: Molecular structure of the rat bone Gla protein gene and identification of promoter  
 A:Reference number: A31419; MUID:89251082; PMID:2765907  
 A:Accession: A31419  
 A:Molecule type: DNA  
 A:Residues: 1-99 <THE>  
 A:Cross-references: GB:M25490; NID:g576530; PIDN:AAA53280.1; PID:g576531  
 R:Lian, J.; Stewart, C.; Puchacz, E.; MacKowalak, S.; Shalhoub, V.; Colvard, D.; Zambae  
 Proc. Natl. Acad. Sci. U.S.A. 86, 1143-1147, 1989  
 A:Title: Structure of the rat osteocalcin gene and regulation of vitamin D-dependent  
 A:Reference number: A32324; MUID:89145200; PMID:2764002  
 A:Accession: A32324  
 A:Molecule type: DNA  
 A:Residues: 1-99 <LIR>  
 A:Cross-references: GB:U04500; NID:g205863; PIDN:AAA41764.1; PID:g205864  
 R:Pan, L.C.; Price, P.A.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 6109-6113, 1985  
 A:Title: The propeptide of rat bone gamma-carboxyglutamic acid protein shares homology  
 A:Reference number: A25167; MUID:85296305; PMID:3875836  
 A:Accession: A25167  
 A:Molecule type: mRNA





A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-3, '1F', 6-10, 'T', 12-95 <RE2>  
A:Cross-references: GB:I24431; NID:g455456; PIDN:AAA39856.1; PID:g455457  
C:Genetics:  
A:Introns: 22/1; 33/1; 52/2  
C:Superfamily: osteocalcin  
F:1-49/Domains: signal sequence #status predicted <SIG>  
F:50-95/Product: osteocalcin #status predicted <MAT>

Query Match 54.4%; Score 148; DB 2; Length 95;  
Best Local Similarity 58.7%; Pred. No. 2.9e-12;  
Matches 27; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 2 LXQWLGAPVPPDPLXPRRXVXCLNPDCELADHIGFOEAYRRFYG 47  
| : ||| || ||| | : ||| ||| : | : ||| ||  
DB 47 LRRYLGASVSPDPLEPTRECELPACDELSDQYGLKAYRRFYG 92

RESULT 10  
153275

osteocalcin - mouse

C:Species: Mus sp. (mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 20-Aug-1999  
C:Accession: I53275

R:Ratman, S.; Oberdorf, A.; Montecino, M.; Tanhauser, S.M.; Llan, J.B.; Stein, G.S.; Lai  
Endocrinology 133, 3050-3053, 1993

A:Title: Multiple copies of the bone-specific osteocalcin gene in mouse and rat.  
A:Reference number: I53275; MUID:94062692; PMID:8243336

A:Accession: I53275

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: GB:S67455; NID:g456854; PIDN:AAB29145.1; PID:g456856

C:Genetics:

A:Introns: 22/1; 33/1; 52/2; 72/2

C:Superfamily: osteocalcin

Query Match 54.4%; Score 148; DB 2; Length 95;  
Best Local Similarity 58.7%; Pred. No. 2.9e-12;  
Matches 27; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 2 LXQWLGAPVPPDPLXPRRXVXCLNPDCELADHIGFOEAYRRFYG 47  
| : ||| || ||| | : ||| ||| : | : ||| ||  
DB 47 LRRYLGASVSPDPLEPTRECELPACDELSDQYGLKAYRRFYG 92

RESULT 11  
167413

osteocalcin - mouse

C:Species: Mus sp. (mouse)

C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Aug-1999  
C:Accession: I67413

R:Ratman, S.; Oberdorf, A.; Montecino, M.; Tanhauser, S.M.; Llan, J.B.; Stein, G.S.; Lai  
Endocrinology 133, 3050-3053, 1993

A:Title: Multiple copies of the bone-specific osteocalcin gene in mouse and rat.  
A:Reference number: I53275; MUID:94062692; PMID:8243336

A:Accession: I67413

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: GB:S67456; NID:g456855; PIDN:AAB29146.1; PID:g456857

C:Genetics:

A:Introns: 22/1; 33/1; 52/2; 72/2

C:Superfamily: osteocalcin

Query Match 52.9%; Score 144; DB 2; Length 95;  
Best Local Similarity 56.5%; Pred. No. 9.7e-12;  
Matches 26; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

OY 2 LXQWLGAPVPPDPLXPRRXVXCLNPDCELADHIGFOEAYRRFYG 47  
| : ||| || ||| | : ||| ||| : | : ||| ||  
DB 47 LRRYLGASVSPDPLEPTRECELPACDELSDQYGLKAYRRFYG 92

RESULT 12  
161188

osteocalcin-related protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 20-Aug-1999  
C:Accession: I61188; I49073

R:Desbois, C.; Hoque, D.A.; Karsenty, G.  
J. Biol. Chem. 269, 1183-1190, 1994

A:Title: The mouse osteocalcin gene cluster contains three genes with two separate sp

A:Reference number: A49871; MUID:94117426; PMID:8288580

A:Accession: I61188

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: GB:I24430; NID:g455454; PIDN:AAA39855.1; PID:g455455

A:Accession: I49073

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RE2>

A:Cross-references: EMBL:U1541; NID:g508297; PIDN:AAB60445.1; PID:g508298

C:Genetics:

A:Introns: 22/1; 33/1; 52/2

C:Superfamily: osteocalcin

Query Match 52.9%; Score 144; DB 2; Length 95;  
Best Local Similarity 56.5%; Pred. No. 9.7e-12;  
Matches 26; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

OY 2 LXQWLGAPVPPDPLXPRRXVXCLNPDCELADHIGFOEAYRRFYG 47  
| : ||| || ||| | : ||| ||| : | : ||| ||  
DB 47 LRRYLGASVSPDPLEPTRECELPACDELSDQYGLKAYRRFYG 92

RESULT 13  
CEMP

osteocalcin - swordfish

N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing prote

C:Species: Xiphias gladius (swordfish)

C:Date: 30-Apr-1979 #sequence\_revision 27-Nov-1985 #text\_change 06-Sep-1996  
C:Accession: A03306

R:Price, P.A.; Otsuka, A.S.; Poser, J.W.  
in Calcium-binding proteins and Calcium function. Wasserman, R.H., Corradino, R.A., C

A:Title: Comparison of gamma-carboxyglutamic acid-containing proteins from bovine and

A:Reference number: A03306

A:Accession: A03306

A:Molecule type: protein

A:Residues: 1-47 <PR1>

A:Note: residues 14, 24, and 37 were not positively identified

C:Superfamily: osteocalcin

C:Keywords: bone; calcium binding; carboxyglutamic acid

F:13-17/20/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:19-25/Disulfide bonds: #status predicted

Query Match 33.5%; Score 91; DB 1; Length 47;  
Best Local Similarity 48.6%; Pred. No. 4.1e-05;  
Matches 18; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

OY 13 PDLXPRRXVXCLNPDCELADHIGFOEAYRRFYGPV 49  
| : ||| || ||| | : ||| ||| : | : ||| ||  
DB 9 PQLLESRLREVCELVNSCEMADTAGIVAYVYGPV 45

RESULT 14  
A42794  
osteocalcin - bluegill  
N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing prote

C:Species: Lepomis macrochirus (bluegill)  
C:Date: 31-Dec-1993 #sequence\_revision 06-Sep-1996 #text\_change 06-Sep-1996  
C:Accession: A42794

R:Nishimoto, S.K.; Araki, N.; Robinson, F.D.; Waite, J.H.  
J. Biol. Chem. 267, 11600-11605, 1992

A>Title: Discovery of bone gamma-carboxyglutamic acid protein in mineralized scales. The

A:Reference number: A42794; MUID:92283881; PMID:1597487

A:Accession: A42794

A:Molecule type: protein

A:Residues: 1-45 <NIS>

A:Experimental source: bone; mineralized scale

A>Note: sequence extracted from NCBI backbone (NCBIF:104759)

C:Superfamily: osteocalcin

C:Keywords: bone; calcium binding; carboxyglutamic acid; vitamin K

F:4/Modified site: gamma-carboxyglutamic acid (Glu) #status absent

F:11,15,18/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental

F:17-23/Dissulfide bonds: #status predicted

Query Match 25.0%; Score 68; DB 1; Length 45;  
Best Local Similarity 40.0%; Pred. No. 0.04; Indels 0; Gaps 0;  
Matches 12; Conservative 5; Mismatches 13;

OY 20 RAVCXILNPDCELDADHIGFOEAYRRFGPV 49  
| | | : : : : | : : : ||  
DB 14 REVCEANLACEDMDAQQGIATVATYGPI 43

RESULT 15

SYEWT

Methionine-tRNA ligase (EC 6.1.1.10) [validated] - Escherichia coli (strain K-12)  
N:Alternate names: methionyl-tRNA synthetase  
C:Species: Escherichia coli  
C:Date: 13-Jun-1993 #sequence\_revision 30-Jun-1992 #text\_change 03-Jun-2002  
C:Accession: S14427; A91797; A91123; S11949; A64979; A01187; A35821  
R:Dardel, F.  
submitted to the EMBL Data Library, October 1990  
A:Reference number: S14427  
A:Accession: S14427

A:Molecule type: DNA

A:Residues: 1-677 <DAR1>

A:Cross-references: EMBL:X55791; NID:g42015; PIDN:CAA93315.1; PID:g42016  
R:Dardel, F.; Fayat, G.; Blanquet, S.  
J. Bacteriol. 160, 1115-1122, 1984

A>Title: Molecular cloning and primary structure of the Escherichia coli methionyl-tRNA  
A:Reference number: A91797; MUID:85054627; PMID:6094501

A:Accession: A91797

A:Molecule type: DNA

A:Residues: 2-677 <DAR2>

A:Cross-references: GB:K02671; NID:g146829; PIDN:AAA24161.1; PID:g146829; GB:J01649; GB:  
R:Barker, D.G.; Edell, J.P.; Jakobs, R.; Bruton, C.J.  
Eur. J. Biochem. 127, 449-457, 1982

A>Title: Methionyl-tRNA synthetase from E. coli: primary structure of the active crystal  
A:Accession: A91123; MUID:83079258; PMID:6756915

A:Molecule type: DNA

A:Residues: 2-15,'V','I','L','I51-435','A','437-564 <BAR>

A:Cross-references: GB:K02671; GB:J01649; GB:J01650

A>Note: most of this sequence was confirmed by protein sequencing

Kol. Gen. Genet. 223, 121-133, 1990

R:Dardel, F.; Panvert, M.; Fayat, G.

A>Title: Transcription and regulation of expression of the Escherichia coli methionyl-tRNA  
A:Reference number: S11948; MUID:91080852; PMID:2259334

A:Accession: S11949

A:Molecule type: DNA

A:Residues: 1-51;661-677 <DAR>

A:Cross-references: EMBL:X55791

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: A64979

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-677 <BLAT>

A:Cross-references: GB:A0000300; GB:U00096; NID:g1788425; PIDN:AAC75175.1; PID:g1788432.  
A:Experimental source: strain K-12, substrain MG1655  
R:Zeilwer, C.; Risler, J.L.; Brunle, S.

```

J.Mol. Biol.55, 63-81, 1982
A.Title: Crystal structure of Escherichia coli methionyl-tRNA synthetase at 2.5 A res
A.Reference number: A92860; MUID:82192427; PMID:7042967
R.Houtonndj, C.; Schmitter, J.M.; Beauvallet, C.; Blanquet, S.
Biochemistry 29, 8190-8198, 1990
A.Title: Mapping of the active site of Escherichia coli methionyl-tRNA synthetase: Id
the 3'-acceptor end.
A.Reference number: A35821; MUID:91084494; PMID:1702021
R.Fournay, D.; Dardel, F.
submitted to the Brookhaven Protein Data Bank, November 1992
A.Reference number: A51298; PDB:1MDA
A.Contents: annotation; conformation by (1)H-NMR, residues 'GS',139-164
R.Fournay, D.; Dardel, F.; Blanquet, S.
J.Mol. Biol. 231, 1078-1089, 1993
A.Title: Methionyl-tRNA synthetase zinc binding domain. Three-dimensional structure a
A.Reference number: A58691; MUID:93294859; PMID:8515466
A.Contents: annotation; conformation by (1)H-NMR
J.Brunie, S.; Zeller, C.; Rislér, J.L.
J.Mol. Biol. 216, 411-424, 1990
A.Title: Crystallographic study at 2.5 angstroms resolution of the interaction of met
A.Reference number: A58692; MUID:91073404; PMID:2254937
A.Contents: annotation; x-ray crystallography, 2.5 angstroms
C.Genetics:
A.Gene: metG
A:Map position: 46 min
C.Function:
A.Description: EC 6.1.1.10 [validated, MUID:83079258]: catalyzes the ligation of meth
A.Pathway: protein biosynthesis
C.Superfamily: methionine-tRNA ligase
C.Keywords: aminocacyl-tRNA synthetase; ATP; homodimer; ligase; metalloprotein; protei
E:2-677/Product: methionine-tRNA ligase #status predicted <MNT>
F:62-66,137-157,335-341,434-438/Region: tRNA 3'-acceptor end binding
F:452-468/Region: anticodon recognition
F:16,22,53,336/Active site: Tyr, His, Asp, Lys #status predicted
F:146,149,159,162/binding site: zinc (Cys) #status experimental
Query Match 21.9%; Score 59.5; DB 1; Length 677;
Best Local Similarity 27.3%; Pred. No. 8.9;
Matches 12; Conservative 7; Mismatches 16; Indels 9; Gaps 1;
QY 1 XLYOWLGAPVPRPDLXPRRYVCXKLNPCDCDLADHIGQEQEYRR 44
Db 250 IFYFWLDAPIG-----MGSEFNKLDNRGGSVSSTDEYWK 284

```

Search completed: December 4, 2002, 15:44:57  
Job time : 19 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: December 4, 2002, 15:45:41 ; Search time 35 Seconds  
(without alignments)  
186.551 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 49

Sequence: 1 YLYQWLGAPEYPPDPLXPR.....DELADHIGFQEAVERFYGPV 49

## Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

## Database :

A\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	93.9	49	14	AAK32937
2	46	93.9	49	18	AAW34266
3	30	61.2	49	18	AAK32936
4	30	61.2	49	18	AAW34264
5	25	51.0	46	22	AAK31919
6	25	51.0	47	22	AAK31917
7	25	51.0	49	18	AAW34263
8	25	51.0	49	18	AAW34265
9	25	51.0	49	20	AAW39811
10	25	51.0	51	12	AAK10146

11	25	51.0	98	19	AAW76094
12	25	51.0	100	12	AAK10147
13	25	51.0	100	23	AAU10687
14	19	38.8	97	13	AAK20043
15	16	32.7	20	13	AAK26407
16	16	32.7	21	11	AAK06630
17	16	32.7	49	16	AAW01681
18	16	32.7	100	23	AAU10688
19	14	28.6	15	11	AAK06631
20	13	26.5	13	22	AAK1920
21	13	26.5	13	23	AAE18389
22	11	22.4	13	18	AAW34268
23	11	22.4	38	8	AAK1282
24	11	22.4	49	23	AAU85975
25	11	22.4	68	23	AAK17841
26	11	22.4	99	19	AAW76097
27	10	20.4	13	18	AAW04605
28	10	20.4	13	21	AAK58430
29	10	20.4	13	22	AAK65668
30	10	20.4	13	22	AAU02089
31	10	20.4	13	22	AAK70172
32	10	20.4	13	23	AAK15409
33	10	20.4	13	23	AAK83453
34	10	20.4	13	23	AAU98372
35	10	20.4	13	23	AAU78634
36	10	20.4	13	23	AAE18828
37	10	20.4	13	23	AAK48056
38	10	20.4	15	21	AAK70693
39	10	20.4	18	20	AAK05366
40	10	20.4	24	22	AAE08623
41	10	20.4	38	23	AAU88451
42	10	20.4	39	23	AAU88442
43	10	20.4	43	18	AAW19857
44	10	20.4	44	23	AAU88438
45	10	20.4	44	23	AAU88447

## ALIGNMENTS

RESULT 1	
AAK32937	standard; peptide: 49 AA.
XX	
AC	AAK32937;
XX	
DT	05-JUL-1993 (first entry)
XX	
DE	17, 21, 24, Gla human osteocalcin peptide.
XX	
KW	Gamma-carboxyglutamic acid; OS.
XX	
OS	Synthetic.
XX	
FH	Key
FT	Misc-difference 17
FT	/label= OTHER
FT	/note= "OTHER" gamma-carboxyglutamic acid"
FT	Misc-difference 21
FT	/label= OTHER
FT	/note= "OTHER" gamma-carboxyglutamic acid"
FT	Misc-difference 24
FT	/label= OTHER
FT	/note= "OTHER" gamma-carboxyglutamic acid"
XX	
XX	JP05032697-A.
PN	
XX	
XX	09-FEB-1993.
PD	
XX	
XX	Bone Gla protein P.
PF	
XX	31-JUL-1991; 91JP-0213251.
XX	
PR	
XX	31-JUL-1991; 91JP-0213251.
XX	

Human osteocalcin  
Human osteocalcin  
Reference sequence  
Fusion protein for  
Osteocalcin AA 1-1  
Human osteocalcin  
Bone Gla protein.  
Polymorphic varian  
Human osteocalcin  
Bone Gla protein P  
Human BGP peptide  
Gla21-osteocalcin  
Peptide from mamma  
Modified osteocalc  
Peptide presentati  
Rat osteocalcin pr  
Osteocalcin 7-19 f  
E tag peptide, SEQ  
E tag used for rec  
E-tag peptide. Sy  
E tag. Synthetic.  
E tag - peptide ta  
E tag. Synthetic.  
Synthetic E tag pe  
Synthetic E tag pe  
E tag used in the  
E-tag sequence. S  
E. coli antibody d  
HBV specific singl  
EHTag peptide. Sy  
Insulin/Insulin-11  
Insulin/Insulin-11  
Vector PUC119EHIS-  
Insulin/Insulin-11  
Insulin/Insulin-11

PA (TEIJIN ) TEIJIN LTD.  
XX WPI; 1993-088665/11.  
XX Synthetic human osteocalcin for standard in determ. of natural  
PT osteocalcin - prepd. by introducing gamma-carboxyglutamic acid  
PT as fluorenyl protected gp.  
XX  
PS Claim 2; Page 2; 10pp; Japanese.  
XX  
CC The synthetic 17, 21, 24, gamma-carboxyglutamic acid form of human  
CC osteocalcin (OS) was produced by introducing protected Gla. The  
CC substance may be produced in high yield and is useful as standard  
CC for the determination of human OS.  
CC See also AAR32936.  
CC  
XX  
SQ Sequence 49 AA;  
XX  
Query Match 93.9%; Score 46; DB 14; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.2e-42;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
OY 1 YLYQWLGAPVPPDPLXPRRXVCXLPNDCDELADHIGFQEAFFRGVPV 49  
1 YLYQWLGAPVPPDPLXPRRXVCXLPNDCDELADHIGFQEAFFRGVPV 49  
DB  
XX  
RESULT 2  
AAR34266  
ID AAR34266 standard; peptide; 49 AA.  
XX  
AC AAW34266;  
XX  
DT 23-APR-1998 (first entry)  
XX  
DE Gla21-osteocalcin peptide #2.  
XX  
KW Gla17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
KM osteoporosis; human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 17  
FT Modified-site /note= "gamma-carboxyglutamic acid"  
FT Modified-site 21  
FT Modified-site /note= "gamma-carboxyglutamic acid"  
FT Modified-site 24  
FT /note= "gamma-carboxyglutamic acid"  
XX  
PN WO9738309-A1.  
XX  
PD 16-OCT-1997.  
XX  
PF 10-APR-1997; 97WO-JP01246.  
XX  
PR 27-FEB-1997; 97JP-0043331.  
PR 10-APR-1996; 96JP-0088608.  
XX  
PA (EISA ) EISAI CO LTD.  
XX  
PI Kimura T, Morimoto S, Sakakibara S;  
XX  
DR WPI; 1997-512875/47.  
XX  
PT Antibody specific for Gla17-osteocalcin, or its fragment - for  
PT diagnosis of bone disorders such as osteoporosis  
XX  
PS Claim 4; Page 18; 28pp; Japanese.  
XX  
CC This sequence represents a human Gla21-osteocalcin peptide. This sequence  
CC is recognised by the antibody of the invention. The antibody of the  
CC invention is an anti-Gla17-osteocalcin antibody or its fragment, which

CC binds to Gla17-osteocalcin, Gla21-osteocalcin or their fragments. The  
CC antibody can be used for the diagnosis of bone related disorders, such as  
CC osteoporosis.  
XX  
SQ Sequence 49 AA;  
XX  
Query Match 93.9%; Score 46; DB 18; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.2e-42;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
OY 1 YLYQWLGAPVPPDPLXPRRXVCXLPNDCDELADHIGFQEAFFRGVPV 49  
1 YLYQWLGAPVPPDPLXPRRXVCXLPNDCDELADHIGFQEAFFRGVPV 49  
DB  
XX  
RESULT 3  
AAR32936  
ID AAR32936 standard; peptide; 49 AA.  
XX  
AC AAR32936;  
XX  
DT 05-JUL-1993 (first entry)  
XX  
DE 21, 24, Gla human osteocalcin peptide.  
XX  
KW Gamma-carboxyglutamic acid; OS.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 21  
FT /label= OTHER  
FT /note= "OTHER= gamma-carboxyglutamic acid"  
FT Misc-difference 24  
FT /label= OTHER  
FT /note= "OTHER= gamma-carboxyglutamic acid"  
XX  
PN JP05032697-A.  
XX  
PD 09-FEB-1993.  
XX  
PF 31-JUL-1991; 91JP-0213251.  
PR 31-JUL-1991; 91JP-0213251.  
XX  
PA (TEIJIN ) TEIJIN LTD.  
XX  
DR WPI; 1993-088665/11.  
XX  
PT Synthetic human osteocalcin for standard in determ. of natural  
PT osteocalcin - prepd. by introducing gamma-carboxyglutamic acid  
PT as fluorenyl protected gp.  
XX  
PS Claim 1; Page 2; 10pp; Japanese.  
XX  
CC The synthetic 21, 24, gamma-carboxyglutamic acid form of human  
CC osteocalcin (OS) was produced by introducing protected Gla. The  
CC substance may be produced in high yield and is useful as standard  
CC for the determination of human OS.  
CC See also AAR32937.  
XX  
SQ Sequence 49 AA;  
XX  
Query Match 61.2%; Score 30; DB 14; Length 49;  
Best Local Similarity 100.0%; Pred. No. 2.8e-25;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
OY 18 PRRXVCXLPNDCDELADHIGFQEAFFRGVPV 49  
18 PRRXVCXLPNDCDELADHIGFQEAFFRGVPV 49  
DB 18 PRRXVCXLPNDCDELADHIGFQEAFFRGVPV 49  
XX  
RESULT 4

```

AAW34264
ID AAW34264 standard; peptide; 49 AA.
XX
AC AAW34264;
XX
DT 23-APR-1998 (first entry)
XX
DE Glu17-osteocalcin peptide #2.
XX
XX Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;
KW osteoporosis; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 21
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site 24
FT /note= "gamma-carboxyglutamic acid"
FT
PN WO9738309-A1.
XX
PD 16-OCT-1997.
XX
PF 10-APR-1997; 97WO-JP01246.
XX
PR 27-FEB-1997; 97JP-0043331.
XX
PR 10-APR-1996; 96JP-0088608.
XX
PA (EISA ) EISAI CO LTD.
XX
PI Kimura T, Morimoto S, Sakakibara S;
DR WPI, 1997-512875/47.
XX
XX
PT Antibody specific for Glu17-osteocalcin, or its fragment - for
PT diagnosis of bone disorders such as osteoporosis
XX
XX Claim 3; Page 16-17; 28pp; Japanese.
XX
XX This sequence represents a human Glu17-osteocalcin peptide. This sequence
CC is recognised by the antibody of the invention. The antibody of the
CC invention is an anti-Glu17-osteocalcin antibody or its fragment, which
CC binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The
CC antibody can be used for the diagnosis of bone related disorders, such as
CC osteoporosis.
XX
SQ Sequence 49 AA;
XX
Query Match 61.2%; Score 30; DB 18; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 18 PRRVXCXLPNDCDELADHIGFOEAYRRFGPV 49
DB 18 PRRVXCXLPNDCDELADHIGFOEAYRRFGPV 49
XX
RESULT 5
AAB91919
ID AAB91919 standard; Peptide; 46 AA.
XX
AC AAB91919;
XX
XX 22-JUN-2001 (first entry)
XX
DE Bone Gla protein peptide SEQ ID NO:1095.
XX
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidy1; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS

```

```

OS Synthetic.
XX
XX WO200069900-A2.
XX
XX 23-NOV-2000.
XX
XX 17-MAY-2000; 2000WO-US13576.
XX
XX 17-MAY-1999; 99US-0134406.
XX
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
XX (CONJ-) CONJUCHEM INC.
XX
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;
PI WPI, 2001-112059/12.
XX
DR
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT
XX
XX Disclosure; Page 553; 733pp; English.
XX
XX The present invention describes a modified therapeutic peptide (I)
XX comprising a therapeutically active amino acid region (III) and a
XX reactive group (II) (e.g. succinimidy1 and maleimido groups) attached to
XX a less therapeutically active amino acid region (IV), which covalently
XX bonds with amino/hydroxyl/thiol groups on blood components to form a
XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
XX (II) are useful for modifying therapeutic peptides e.g. hormones, growth
XX factors and neurotransmitters, to protect them from peptidase activity
XX in vivo for the treatment of various disorders. Endogenous therapeutic
XX peptides are not suitable as drug candidates as they require frequent
XX administration due to rapid degradation by peptidases in the body.
XX Modifying and attaching therapeutic peptides to albumin prevents or
XX reduces the action of peptidases to increase length of activity (half
XX life) and specificity as bonding to large molecules decreases
XX intracellular uptake and interference with physiological processes.
XX AAB90829 to AAB92441 represent peptides which can be used in the
XX exemplification of the present invention.
XX
SQ Sequence 46 AA;
XX
Query Match 51.0%; Score 25; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 7.1e-20;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 25 LNPDCDELADHIGFOEAYRRFGPV 49
DB 22 LNPDCDELADHIGFOEAYRRFGPV 46
XX
RESULT 6
AAB91917
ID AAB91917 standard; Peptide; 47 AA.
XX
AC AAB91917;
XX
XX 22-JUN-2001 (first entry)
XX
DE Bone Gla protein peptide SEQ ID NO:1093.
XX
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidy1; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO200069900-A2.
XX
XX 23-NOV-2000.
XX

```

XX 17-MAY-2000; 2000WO-US13576.  
PF  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX  
DR WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
PS Disclosure; Page 552-553; 733pp; English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide) and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specifically as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 47 AA:  
Query Match 51.0%; Score 25; DB 22; Length 47;  
Best Local Similarity 100.0%; Pred. No. 7.3e-20;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
DB 23 LNPDCDELADHIGFOEAYRRFGPV 47  
RESULT 7  
AAW34263  
ID AAW34263 standard; peptide: 49 AA.  
XX  
AC AAW34263;  
XX  
DT 23-APR-1998 (first entry)  
XX  
DE Glu17-osteocalcin peptide #1.  
XX  
KW Glu17-osteocalcin; Glu21-osteocalcin; antibody; bone disorder; diagnosis;  
KW osteoporosis; human.  
XX  
OS Homo sapiens.  
XX  
FT Key Location/Qualifiers  
FT Modified-site 21 /note= "gamma-carboxyglutamic acid"  
XX  
FN W09738309-A1.  
XX  
PD 16-OCT-1997.  
XX  
PF 10-APR-1997; 97WO-JP01246.  
XX

PR 27-FEB-1997; 97JP-0043331.  
PR 10-APR-1996; 96JP-0088608.  
XX  
PA (EISA ) EISAI CO LTD.  
XX  
PI Kimura T, Morimoto S, Sakakibara S;  
XX  
DR WPI; 1997-512875/47.  
XX  
PT Antibody specific for Glu17-osteocalcin, or its fragment - for  
PT diagnosis of bone disorders such as osteoporosis  
XX  
PS Claim 3; Page 16; 28pp; Japanese.  
XX  
CC This sequence represents a human Glu17-osteocalcin peptide. This sequence  
CC is recognised by the antibody of the invention. The antibody of the  
CC invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
CC binds to Glu17-osteocalcin, Glu21-osteocalcin or their fragments. The  
CC antibody can be used for the diagnosis of bone related disorders, such as  
CC osteoporosis.  
XX  
SQ Sequence 49 AA:  
Query Match 51.0%; Score 25; DB 18; Length 49;  
Best Local Similarity 100.0%; Pred. No. 7.5e-20;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
DB 25 LNPDCDELADHIGFOEAYRRFGPV 49  
RESULT 8  
AAW34265  
ID AAW34265 standard; peptide: 49 AA.  
XX  
AC AAW34265;  
XX  
DT 23-APR-1998 (first entry)  
XX  
DE Glu21-osteocalcin peptide #1.  
XX  
KW Glu17-osteocalcin; Glu21-osteocalcin; antibody; bone disorder; diagnosis;  
KW osteoporosis; human.  
XX  
OS Homo sapiens.  
XX  
FT Key Location/Qualifiers  
FT Modified-site 17 /note= "gamma-carboxyglutamic acid"  
FT Modified-site 21 /note= "gamma-carboxyglutamic acid"  
XX  
FN W09738309-A1.  
XX  
PD 16-OCT-1997.  
XX  
PE 10-APR-1997; 97WO-JP01246.  
XX  
PR 27-FEB-1997; 97JP-0043331.  
PR 10-APR-1996; 96JP-0088608.  
XX  
PA (EISA ) EISAI CO LTD.  
XX  
PI Kimura T, Morimoto S, Sakakibara S;  
XX  
DR WPI; 1997-512875/47.  
XX  
PT Antibody specific for Glu17-osteocalcin, or its fragment - for  
PT diagnosis of bone disorders such as osteoporosis  
XX  
PS Claim 4; Page 17; 28pp; Japanese.  
XX



CC This sequence represents a human Gla21-osteocalcin peptide. This sequence  
CC is recognised by the antibody of the invention. The antibody of the  
CC invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
CC binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The  
CC antibody can be used for the diagnosis of bone related disorders, such as  
CC osteoporosis.

XX Sequence 49 AA;

Query Match 51.0%; Score 25; DB 18; Length 49;  
Best Local Similarity 100.0%; Pred. No. 7.5e-20;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
Db 25 LNPDCDELADHIGFOEAYRRFGPV 49  
|||||

RESULT 9  
AAM99811

ID AAM99811 standard; Protein: 49 AA.

AC AAM99811;

DT 08-JUN-1999 (first entry)

DE Human osteocalcin fragment.

KW Human; osteocalcin; urine; gamma-carboxylated; bone disorder; hOC;

KM bone turnover; puberty; menopause; detection.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Modified-site 17 /note- "possibly gamma-carboxylated"

FT Modified-site 21 /note- "possibly gamma-carboxylated"

FT Modified-site 24 /note- "possibly gamma-carboxylated"

PN WO9909058-A1.

PD 25-FEB-1999.

PE 24-JUN-1998; 98WO-FI00550.

PR 15-AUG-1997; 97FI-0003371.

PA (HELL/) HELLMAN J.

PA (KAER/) KAERKONEN S.

PA (KARP/) KARP M.

PA (LOEV/) LOEVGREN T.

PA (PETT/) PETTERSSON K.

PA (VAEA/) VAEANEN H K.

PI Hellman J, Kaerkoenen S, Karp M, Loevgren T, Pettersson K;

PI Vaeanaenen HK;

DR WPI: 1999-180972/15.

DR N-PSDB: AAX19767.

PT Human osteocalcin peptide fragments - useful for detecting

PT metabolic bone turnover rates and disorders

PS Claim 1; Fig 1A: 49pp; English.

XX An assay has been developed for the detection of gamma-carboxylated  
CC human osteocalcin found in urine, where the glutamic acid residues at  
CC positions 17, 21 and 24 are gamma-carboxylated. The assay is used to  
CC measure the rate of bone turnover (formation and/or resorption) and/or  
CC for investigating metabolic bone disorders in individuals. Especially  
CC mentioned, the assays can detect differences between adults and children

CC going through puberty, pre- and post-menopausal women and children  
CC having high bone turnover. The assay allows more sensitive detection of  
CC human osteocalcin (hOC) in urine than in serum. Significant increases in  
CC hOC (40-48%) were detected in serum in menopause (normal increase is  
CC 30-50% above pre-menopausal women). Detection of urine hOC using the  
CC assay showed an increase as high as 75-79%. The present sequence  
CC represents a fragment of hOC.

XX Sequence 49 AA;

Query Match 51.0%; Score 25; DB 20; Length 49;  
Best Local Similarity 100.0%; Pred. No. 7.5e-20;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
Db 25 LNPDCDELADHIGFOEAYRRFGPV 49  
|||||

RESULT 10  
AAR10146

ID AAR10146 standard; peptide: 51 AA.

AC AAR10146;

DT 27-MAR-1991 (first entry)

DE Human osteocarcin precursor polypeptide (I).

KW Human osteocarcin precursor polypeptide; OC; carboxypeptidase B;

KM calcium; vitamin K; bone formation; dysbolism.

OS Homo sapiens.

FT Key Location/Qualifiers

FT MISC-difference 1..1 /note- "residue may be omitted"

FT MISC-difference 51..51 /note- "residue may be omitted"

PN JP02201294-A.

PD 03-DEC-1990.

PE 22-AUG-1989; 89JP-0214239.

PR 06-DEC-1988; 88JP-0306931.

PR 22-AUG-1989; 89JP-0214239.

PA (TAKA-) TAKARA SYUZO KK.

DR WPI: 1991-018865/03.

PT Human osteocarcin (OC) precursor polypeptide - used to prepare

PT purified OC by digestion with carboxypeptidase B.

PS Claim 1; Page 1; 11pp; Japanese.

XX To K1 is attached H and to K51 is attached OH.

CC A novel gene encoding human OC precursor polymer was inserted into

CC plasmid pOC 980. E.coli HB101 was transformed with this plasmid, and

CC made to efficiently produce the objective polymer. The polymer was

CC treated with lysylendopeptidase B to remove spacers. The obtained

CC human precursor was further treated with carboxy peptidase to prepare a

CC purified human OC.

CC OC is a calcium bonded protein (depending on vitamin K) produced in the

CC bone. It is thought to be a promoting factor during the bone

CC formation and may be used to treat diseases due to dysbolism of the

CC bone.  
See also AAR10147 and AAQ10193-98.

XX Sequence 51 AA;

Query Match 51.0%; Score 25; DB 12; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-20;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDCELADHIGFOEAYRRFGPV 49  
 DB 26 LNPDCELADHIGFOEAYRRFGPV 50

RESULT 11  
 AAW76094  
 ID AAW76094 standard; Protein; 98 AA.

XX AAW76094;

DT 21-DEC-1998 (first entry)

XX Human osteocalcin protein.

XX Promoter; tissue-specific gene expression; skeletal tissue;

KW stem cell; bone; cartilage; osteocalcin; hOC gene; human;

KW osteoporosis; osteopenia; osteosarcoma; cancer; metastasis;

XX gene therapy.

OS Homo sapiens.

XX WO9839427-A2.

XX 11-SEP-1998.

XX 06-MAR-1998; 98WO-US04421.

XX 06-MAR-1997; 97US-0039839.

XX (UYMA-) UNIV MASSACHUSETTS.

XX Frenkel B, Hou Z, Lian JB, Nilsson S, Peters S;

PI Quesenberry P, Stein GS, Stein JL;

DR N-PSDB; AAV46429.

XX WPI: 1998-495839/42.

XX Expression of exogenous genes in differentiated cells - by

PT transducing pluripotent stem cells capable of maturing into

PT differentiated cells with nucleic acid comprising exogenous gene,

XX useful for, e.g. treatment of osteoporosis

PS Disclosure; Page 33-34; 63pp; English.

XX This is the amino acid sequence of the human osteocalcin, deduced

CC from the coding exons of the human hOC gene (see AAV46429). The

CC invention pertains to a method for expressing endogenous genes in

CC differentiated cells of a specific type. The method involves

CC contacting pluripotent stem cells capable of maturing into

CC differentiated cells with a nucleic acid comprising an exogenous

CC gene linked to a regulatory element capable of controlling expression

CC of the exogenous gene in the differentiated cells. A population of

CC transduced stem cells capable of maturing into differentiated cells

CC expressing the exogenous gene is produced. Preferably, the

CC differentiated cells are in a tissue of interest, such as bone or

CC cartilage, and the exogenous gene is operably linked to at least one

CC osteocalcin regulatory element derived from the hOC promoter. The

CC exogenous gene can encode a therapeutic protein useful for treating

CC a disease, especially osteoporosis, osteopenia, osteosarcoma,

CC primary malignancy or metastases (all claimed).

XX Sequence 98 AA;

XX Query Match 51.0%; Score 25; DB 19; Length 98;

XX Best Local Similarity 100.0%; Pred. No. 1.4e-19;

XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 25 LNPDCELADHIGFOEAYRRFGPV 49

DB 74 LNPDCELADHIGFOEAYRRFGPV 98

RESULT 12  
 AAR10147  
 ID AAR10147 standard; Protein; 100 AA.

XX AAR10147;

DT 27-MAR-1991 (first entry)

XX Human osteocalcin precursor polypeptide (II).

XX Human osteocalcin precursor polypeptide; OC; carboxypeptidase B;

KW calcium; vitamin K; bone formation; dysbolism.

XX Homo sapiens.

XX JP02201294-A.

XX 03-DEC-1990.

XX 22-AUG-1989; 89JP-0214239.

XX 06-DEC-1988; 88JP-0306931.

XX 22-AUG-1989; 89JP-0214239.

XX (TAKA-) TAKARA SYUZO KK.

XX WPI: 1991-01865/03.

XX Human osteocalcin (OC) precursor polypeptide - used to prepare

PT purified OC by digestion with carboxypeptidase B.

XX Claim 5; Page 1; 11pp; Japanese.

XX To K1 is attached H and to V100 is attached OH.

CC A novel gene encoding human OC precursor polymer was inserted into

CC plasmid pOC 980. E.coli HB101 was transformed with this plasmid, and

CC made to efficiently produce the objective polymer. The polymer was

CC treated with lysylendopeptidase B to remove spacers. The obtained

CC human precursor was further treated with carboxypeptidase to prepare a

CC purified human OC.

CC OC is a calcium bonded protein (depending on vitamin K) produced in the

CC bone. It is thought to be a promoting factor during the bone

CC formation and may be used to treat diseases due to dysbolism of the

CC bone.

CC See also AAR10147 and AAQ10193-98.

XX Sequence 100 AA;

XX Query Match 51.0%; Score 25; DB 12; Length 100;

XX Best Local Similarity 100.0%; Pred. No. 1.4e-19;

XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 25 LNPDCELADHIGFOEAYRRFGPV 49

DB 76 LNPDCELADHIGFOEAYRRFGPV 100

RESULT 13  
 AAU10687  
 ID AAU10687 standard; Protein; 100 AA.

XX AAU10687;

XX 14-FEB-2002 (first entry)

XX Reference sequence for human BGLAP protein.

XX Human; single nucleotide polymorphism; SNP; BGLAP; chromosome 1q25-q31;

XX bone gamma carboxyglutamate protein; haplotyping; genotyping;

KW

KW	osteoporosis; osteopathia.
XX	
OS	Homo sapiens.
PN	WO200177131-A2.
XX	
PD	18-OCT-2001.
XX	
PF	11-APR-2001; 2001WO-US12010.
XX	
PR	11-APR-2000; 2000US-195840P.
XX	
PA	(GENA-) GENAISSANCE PHARM INC.
XX	
PI	Bentivegna SC, Chew A, Choi JY, Koshy B, Rounds E, Stephens JC;
XX	
DR	WPI: 2002-041288/05.
DR	N-PSTDB: AAS16363, AAS16364.
XX	
PT	New haplotypes of the human bone gamma carboxyglutamate protein gene,
PR	useful to diagnose and treat diseases associated with the gene such as
XX	osteoporosis
PS	Clam 27; Fig 3; 53pp; English.
XX	
CC	The present invention relates to novel single nucleotide polymorphisms
CC	(SNPs) in the human bone gamma carboxyglutamate protein (BGLAP) gene
CC	located on chromosome 1q25-q31, and methods for haplotyping and/or
CC	genotyping the BGLAP gene in an individual. The methods of the
CC	invention make use of allele-specific oligonucleotides (ASOs) as probes
CC	and primers and/or primer-extensions oligonucleotides for detecting the
CC	BGLAP gene polymorphisms. The polymucleotides and screened compounds are
CC	useful for (developing) treatment of diseases associated with BGLAP
CC	activity, such as osteoporosis. The present sequence represents a
CC	reference sequence for the BGLAP protein.
XX	
SO	Sequence 100 AA:
	Query Match 51.0%; Score 25; DB 23; Length 100;
	Best Local Similarity 100.0%; Pred. No. 1,4e-19;
	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	25 LNPDCDELADHIGFQEAAYRREYGPV 49
DB	76 LNPDCDELADHIGFQEAAYRREYGPV 100
RESULT 14	
AAAR20043	
ID	AAAR20043 standard; Protein; 97 AA.
XX	
AC	AAAR20043;
XX	
DT	09-APR-1992 (first entry)
XX	
DE	Fusion protein for expression of human osteocalcin.
XX	
KW	gamma-carboxyglutamic acid; bone matrix; Gla protein.
XX	
OS	Synthetic.
XX	
OS	
FH	Key
FT	Peptide
FT	Peptide
FT	Cleavage-site
FT	Protein
XX	
PN	EP463571-A.
XX	
PD	02-JAN-1992.

XX	20-JUN-1991;	91EP-0110173.
PF		
XX	30-NOV-1990;	90JP-0330146.
PR	20-JUN-1990;	90JP-0159909.
XX		
PA	(TEIJ ) TEIJIN KK.	
XX		
PI	Eguchi H, Kamimura TF, Sugiyama T, Hosoda K;	
DR	WPI; 1992-009183/02.	
XX	N-PSDB; AAO20210.	
XX		
PT	Human osteocalcin prodn. - using DNA coding for human osteocalcin	
XX	fusion protein for expression in host cells	
PS	Claim 15; Fig 7; 53pp; English.	
XX		
CC	This sequence is a specific example of a claimed generic fusion	
CC	protein comprising human osteocalcin. The pro-peptide is recognised	
CC	by an enzyme capable of Glu to Gla conversion on human osteocalcin.	
CC	The recombinant protein was obtained by culturing host cells	
CC	transformed with a vector containing the synthetic coding sequence.	
CC	The Glu residues could then be converted to Gla (i.e.	
CC	gamma-carboxyglutamic acid) and the osteocalcin sequence cleaved	
CC	from the propeptide. The mature protein is suitable for use in	
CC	immunoassays and as a drug for treatment of bone metabolism	
CC	disorders. See also AAR20044-6.	
XX		
SO	Sequence	97 AA;
	Query Match	38.8%; Score 19; DB 13; Length 97;
	Best Local Similarity	100.0%; Pred. No. 4.4e-13;
	Matches 19; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	25 LNPDCDELADHIGFOEAYR 43	
DB	73 LNPDCDELADHIGFOEAYR 91	
	RESULT 15	
AAR26407		
ID	AAR26407 standard; peptide; 20 AA.	
XX		
AC	AAR26407;	
XX		
DT	20-MAY-1998 (first entry)	
XX		
DE	Osteocalcin AA 1-19, peptide Abi.	
XX		
KW	Osteocalcin; antibody; bone; absorption; metabolism.	
XX		
OS	Synthetic.	
PN	JP04225162-A.	
XX		
PD	14-AUG-1992.	
XX		
PF	27-DEC-1990; 90JP-0415242.	
XX		
PR	27-DEC-1990; 90JP-0415242.	
XX		
PA	(TEIJ ) TEIJIN LTD.	
XX		
DR	WPI; 1992-320370/39.	
XX		
PT	Antibody to human osteocalcin - obt'd. by immunising with	
XX	peptide of 1 to 19th radicals of N-terminal of human	
XX	osteocalcin	
XX		
XX	Disclosure: Fig 1; 8pp; Japanese.	
XX		
XX	The sequences given in AAR26407-9 correspond to fragments of human	

CC osteocalcin. These fragments were used in the production of an  
 CC antibody to human osteocalcin. The fragments were injected into an  
 CC animal and the resulting antibodies were obtained. The antibodies  
 CC can be used in the determination of the rate of bone absorption in  
 CC bone metabolism.

XX Sequence 20 AA;

Query Match 32.7%; Score 16; DB 13; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLXQWLGAAPVPPDPL 16  
 |||||||  
 DB 2 YLXQWLGAAPVPPDPL 17

RESULT 16  
 AAR06630  
 ID AAR06630 standard; peptide; 21 AA.

XX AAR06630;

DT 09-JAN-1991 (first entry)

XX Human osteocalcin fragment.

XX Osteocalcin; musculoskeletal disorders; sandwich assay.

XX WO9009587-A.

XX 23-AUG-1990.

XX 08-FEB-1990; 90WO-JP00155.

XX 02-OCT-1989; 89JP-0255306.

PR 10-FEB-1989; 89JP-0030003.

XX (TEIJU) TEIJU KK.

PI Hosoda K, Honda H, Kubota T, Masuho Y;

XX WPI; 1990-275233/36.

PT Immunassay of human osteocalcin in diagnostic specimens - using  
 PT sandwich assay with antibodies raised respectively to N-terminal  
 PT and C-terminal sites on osteocalcin

XX Disclosure; fig 1; 80pp; Japanese.

XX A monoclonal antibody (Mab), OST-N20, is raised to this N-terminal  
 CC peptide (bases 1-20) of human osteocalcin. Additional Mabs (OST-  
 CC C7 and OST-C15) are raised to two C-terminal osteocalcin peptide  
 CC fragments (comprising bases 43-49 and 35-49 respectively). A  
 CC sandwich immunoassay is then carried out and musculoskeletal dis-  
 CC orders can be diagnosed. See also AAR06631.

XX Sequence 21 AA;

Query Match 32.7%; Score 16; DB 11; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLXQWLGAAPVPPDPL 16  
 |||||||  
 DB 1 YLXQWLGAAPVPPDPL 16

RESULT 17  
 AAW01681  
 ID AAW01681 standard; protein; 49 AA.  
 XX  
 AC AAW01681;

XX 01-APR-1997 (first entry)  
 DT Bone Gla protein.

XX BGP; bone gla protein; osteocalcin; Vitamin K-dependent protein;  
 KW bone matrix; therapy; diagnosis; assay; metabolic bone disease.

XX Homo sapiens.

OS Key Location/Qualifiers

FH Misc-difference 1 /note= "underlined in specification"

FT Misc-difference 3 /note= "underlined in specification"

FT Misc-difference 12 /note= "underlined in specification"

FT Misc-difference 21 /note= "underlined in specification"

FT Modified-site 21 /label= OTHER

FT /note= "gamma-carboxyglutamic acid"

FT Misc-difference 23..29 /note= "underlined in specification"

FT Modified-site 24 /label= OTHER

FT /note= "gamma-carboxyglutamic acid"

FT Misc-difference 42 /note= "underlined in specification"

FT Misc-difference 46 /note= "underlined in specification"

XX DE4340597-AI.

XX 01-JUN-1995.

XX 29-NOV-1993; 93DE-4340597.

XX 29-NOV-1993; 93DE-4340597.

XX (HENN-) HENNING BERLIN GMBH.

PA (BRAH-) BRAHMS DIAGNOSTICA GMBH.

XX Bergmann A, Weckermann R;

PI WPI; 1995-201516/27.

XX Determn. of osteocalcin in serum or plasma - with addn. of divalent  
 PT metal ions to inhibit decompn. of the protein, useful in therapy  
 PT and diagnosis of bone disease

XX Disclosure; Column 1; 9pp; German.

XX A method for determining concentration of osteocalcin in serum or plasma  
 CC is improved with addition of divalent metal ions to inhibit decomposition  
 CC of the protein. The method is useful in therapy and diagnosis of bone  
 CC disease. The present sequence is osteocalcin (a vitamin K-dependent  
 CC protein that is a component of the bone matrix, or alternatively bone gla  
 CC protein).

XX Sequence 49 AA;

Query Match 32.7%; Score 16; DB 16; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-10;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLXQWLGAAPVPPDPL 16  
 |||||||  
 DB 1 YLXQWLGAAPVPPDPL 16

RESULT 18  
 AAU10688  
 ID AAU10688 standard; Protein; 100 AA.  
 XX

```

AC  AAU10688:
XX
DT  14-FEB-2002 (first entry)
XX
DE  Polymorphic variant of human BGLAP protein.
XX
KM  Human; single nucleotide polymorphism: SNP: BGLAP; chromosome 1q25-q31;
KW  bone gamma carboxyglutamate protein; haplotyping; genotyping;
KM  osteoporosis; osteopathic; variant.
XX
OS  Homo sapiens.
XX
FH  Key
FT  Misc-difference 82 Location/Qualifiers
FT  Misc-difference 94 /note="Substitution of Glu to Lys"
FT  Misc-difference 94 /note="Substitution of Arg to Gln"
XX
PN  WO200177131-A2.
XX
PD  18-OCT-2001.
XX
PF  11-APR-2001; 2001WO-US12010.
XX
PR  11-APR-2000; 2000US-195840P.
XX
PA  (GENA-) GENAISSANCE PHARM INC.
XX
PI  Bentivegna SC, Chew A, Choi JY, Koshy B, Rounds E, Stephens JC;
XX  MPI; 2002-041288/05.
XX
DR  New haplotypes of the human bone gamma carboxyglutamate protein gene,
PT  useful to diagnose and treat diseases associated with the gene such as
PT  osteoporosis.
XX
PS  Claim 27; Page -: 53pp; English.
XX
CC  The present invention relates to novel single nucleotide polymorphisms
CC  (SNPs) in the human bone gamma carboxyglutamate protein (BGLAP) gene
CC  located on chromosome 1q25-q31, and methods for haplotyping and/or
CC  genotyping the BGLAP gene in an individual. The methods of the
CC  invention make use of allele-specific oligonucleotides (ASOs) as probes
CC  and primers and/or primer-extension oligonucleotides for detecting the
CC  BGLAP gene polymorphisms. The polynucleotides and screened compounds are
CC  useful for (developing) treatment of diseases associated with BGLAP
CC  activity, such as osteoporosis. The present sequence represents a
CC  polymorphic variant of the BGLAP protein (AAU10687).
CC  Note: The present sequence is not given in the specification but is
CC  created by the indexer from the information given in the patent.
XX
SQ  Sequence 100 AA:
XX
Query Match 32.7%; Score 16; DB 23; Length 100;
Best Local Similarity 100.0%; Pred. No. 8.3e-10;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  1 YLYQWIGAPVPYPDPL 16
    |||||||||||||||
DB  52 YLYQWIGAPVPYPDPL 67

```

```

XX
FH  Key Location/Qualifiers
FT  Peptide 9..15
XX
PN  WO9009587-A.
XX
PD  23-AUG-1990.
XX
PF  08-FEB-1990; 90WO-JP00155.
XX
PR  02-OCT-1989; 89JP-0255306.
XX  10-FEB-1989; 89JP-0030003.
XX
PA  (TEIJ ) TEIJIN KK.
XX
PI  Hosoda K, Honda H, Kubota T, Masuho Y;
XX  MPI; 1990-275233/36.
XX
DR  Immunassay of human osteocalcin in diagnostic specimens - using
PT  sandwich assay with antibodies raised respectively to N-terminal
PT  and C-terminal sites on osteocalcin
XX
PS  Disclosure; fig 3; 80pp; Japanese.
XX
CC  Monoclonal antibodies (Mabs), OST-C7 and OST-CL5 are raised to
CC  this osteocalcin peptide fragment (C-terminal bases 43-49) and
CC  a shorter constituent of this (C-terminal bases 35-49) resp-
CC  ectively. A further MAb (OST-N20) is raised to a fragment
CC  comprising N-terminal bases 1-20. A sandwich immunassay is
CC  then carried out and musculoskeletal dis- orders can be diag-
CC  nosed. See also AAR06630.
XX
SQ  Sequence 15 AA:
XX
Query Match 28.6%; Score 14; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  36 IGFOEAYRREYGPV 49
    |||||||||||||||
DB  2 IGFOEAYRREYGPV 15

```

```

RESULT 20
AAB91920
ID  AAB91920 standard; Peptide; 13 AA.
XX
AC  AAB91920;
XX
DT  22-JUN-2001 (first entry)
XX
DE  Bone Gla protein peptide SEQ ID NO:1096.
XX
KW  Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW  blood component; modification; succinimidyl; maleimido group; amino;
KW  hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS  Homo sapiens.
XX  Synthetic.
XX
PN  WO200069900-A2.
XX
PD  23-NOV-2000.
XX
PF  17-MAY-2000; 2000WO-US13576.
XX
PR  17-MAY-1999; 99US-0134406.
XX  10-SEP-1999; 99US-0153406.
XX  15-OCT-1999; 99US-0159783.
XX
PA  (CONJ-) CONJUCHEM INC.
XX

```

AAR06631  
 ID AAR06631 standard; protein; 15 AA.  
 AC AAR06631;  
 DT 09-JAN-1991 (first entry)  
 DE Human osteocalcin fragment.  
 KW Osteocalcin; musculoskeletal disorders; sandwich assay.

PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
DR WPI: 2001-112059/12.  
XX  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PS  
XX Disclosure: Page 554; 733pp; English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimido) and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 13 AA:  
Query Match 26.5%; Score 13; DB 22; Length 13;  
Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 37 GFOEAYRRFGPV 49  
DB 1 GFOEAYRRFGPV 13  
RESULT 21  
AAE18389  
ID AAE18389 standard; peptide; 13 AA.  
XX  
AC AAE18389;  
XX  
DT 07-MAY-2002 (first entry)  
XX  
DE Human BGP peptide #1.  
XX  
KW Human; cell proliferation; cell differentiation; parathyroid hormone;  
KW PTH; parathyroid related peptide; hyperproliferative skin disorder;  
KW psoriasis; ichthyosis; actinic keratosis; alopecia; skin cell growth;  
KW hair growth; wrinkle; wound healing; chemotherapy; skin cancer; PTHRP;  
KW bone Gla protein; BGP.  
XX  
OS Homo sapiens.  
XX  
PN WO200198348-A2.  
XX  
PD 27-DEC-2001.  
XX  
PE 20-JUN-2001; 2001WO-US19650.  
XX  
PR 22-JUN-2000; 2000US-213247P.  
XX  
PA (HOLI/) HOLICK M F.  
PI Holick MF;  
XX WPI: 2002-171552/22.  
DR Modulating proliferation or differentiation of mammalian skin or hair  
PT cell for treating hyperproliferative diseases, comprises topical

PT administration of liposome encapsulated-peptide identical to  
PT parathyroid hormone -  
XX  
XX  
PS Disclosure: Page 10; 58pp; English.  
XX  
CC The invention relates to a method of modulating proliferation or  
CC differentiation of mammalian skin or hair cell. The method involves  
CC topical administration of a liposome encapsulated-peptide, its salt  
CC or derivative, identical to parathyroid hormone (PTH) or parathyroid  
CC related peptide (PTHrP). Method of the invention is useful for  
CC inhibiting hyperproliferative skin disorders e.g. psoriasis,  
CC ichthyosis, actinic keratosis, skin cancer, inhibiting hair growth or  
CC preventing hair regrowth. It is also useful for stimulating skin cell  
CC growth, rejuvenating aged skin, preventing and treating skin wrinkles,  
CC enhancing wound healing, stimulating hair growth, maintaining hair  
CC growth, treating or preventing female or male pattern baldness,  
CC treating chemotherapy induced alopecia and for stimulating epidermal  
CC cell growth and hair follicle cell growth. The present sequence is  
CC a peptide of human bone Gla protein (BGP) used in the method of the  
CC invention.  
XX  
SQ Sequence 13 AA:  
Query Match 26.5%; Score 13; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 37 GFOEAYRRFGPV 49  
DB 1 GFOEAYRRFGPV 13  
RESULT 22  
AAW34268  
ID AAW34268 standard; peptide; 13 AA.  
XX  
AC AAW34268;  
XX  
DT 23-APR-1998 (first entry)  
XX  
DE Gla21-osteocalcin peptide fragment.  
XX  
KW Gla17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
KW osteoporosis; human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 8  
FT Modified-site 12  
FT Modified-site /note="gamma-carboxyglutamic acid"  
FT Modified-site /note="gamma-carboxyglutamic acid"  
XX  
PN WO9738309-A1.  
XX  
PD 16-OCT-1997.  
XX  
PE 10-APR-1997; 97WO-JP01246.  
XX  
PR 27-FEB-1997; 97JP-0043331.  
XX  
PR 10-APR-1996; 96JP-0088608.  
XX  
PA (EISA ) EISAI CO LTD.  
XX  
PI Kimura T, Morimoto S, Sakakibara S;  
XX WPI: 1997-512875/47.  
DR Antibody specific for Glu17-osteocalcin, or its fragment - for  
PT diagnosis of bone disorders such as osteoporosis  
PS Claim 6; Page 18; 28pp; Japanese.  
XX

CC This sequence represents fragment of a human Gli21-osteocalcin peptide  
(see AAM34265). This sequence is recognised by the antibody of the  
CC invention. The antibody of the invention is an anti-Gli21-osteocalcin  
CC antibody or its fragment, which binds to Gli21-osteocalcin,  
CC Gli21-osteocalcin or their fragments. The antibody can be used for the  
CC diagnosis of bone related disorders, such as osteoporosis.

XX Sequence 13 AA:

Query Match 22.4%; Score 11; DB 18; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.9e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 VPPDPLXPRXV 22  
Db 1 VPPDPLXPRXV 13

RESULT 23

AAP71282  
ID AAP71282 standard; peptide: 38 AA.

XX AAP71282;

DT 11-MAR-1991 (first entry)

DE Peptide from mammalian bone extract.

XX Osteocalcin.

XX Canis familiaris.

OS Location/Qualifiers

FT Modified-site 17

FT /label= G1a /note= "G1a= gamma-carboxylutamic acid"

FT Modified-site 21

FT /label= G1a /note= "G1a= gamma-carboxylutamic acid"

FT Modified-site 24

FT /label= G1a /note= "G1a= gamma-carboxylutamic acid"

XX JPE1263999-A.

XX 21-NOV-1986.

XX 20-MAY-1985; 85JP-0107884.

XX 20-MAY-1985; 85JP-0107884.

XX (TOXN ) TOYO JOZO KK.

XX WPI; 1987-003756/01.

XX New polypeptide used for studying bone metabolism etc. - obd.

XX eg by extracting suspension of bone-meal.

XX Claim 1; Page 967; 4pp; Japanese.

XX Peptide is useful in the study of bone metabolism and estimation of

XX osteocalcin level in the blood.

XX Sequence 38 AA:

Oy 6 LGAPVPPDPL 16  
Db 6 LGAPVPPDPL 16

RESULT 24

AAU85975  
ID AAU85975 standard; protein: 49 AA.

XX AAU85975;

DT 21-MAY-2002 (first entry)

DE Modified osteocalcin peptide.

XX Increased biological potency; prolonged activity; increased half-life;

XX glucose intolerance; insulin resistance; type II diabetes; bone disease;

XX cancer; inflammatory disorder; obesity; developmental disorder;

XX hyperproliferative skin disease; hormone-dependent disease; homeostasis;

XX intestinal disease; interleukin-8 production; smooth muscle contraction;

XX feeding; blood pressure; pancreatic secretion; mutant; mutein;

XX osteocalcin.

OS Unidentified.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 3 /label= Tyr, Asp, Asn

FT Misc-difference 4 /label= Gln, His, Asn

FT Misc-difference 5 /label= Trp, Gly

FT Misc-difference 10 /label= Val, Ala

FT Misc-difference 19 /label= Arg, Lys, His

FT Misc-difference 28 /label= Asp, Asn

FT Misc-difference 40 /label= Glu, Asp

FT Misc-difference 43 /label= Arg, Lys

FT Misc-difference 45 /label= Phe, Ile

FT Misc-difference 48 /label= Pro, Thr

FT Modified-site 49 /note= "C-terminal amide"

XX WO200210195-A2.

XX 07-FEB-2002.

XX 02-AUG-2001; 2001WO-CA01119.

XX 02-AUG-2000; 2000US-222619P.

XX (THER-) THERATECHNOLOGIES INC.

XX Gravel D, Habi A, Abridat T;

XX WPI; 2002-206179/26.

XX Novel modified biological peptide with increased biological potency,

XX prolonged activity, increased half-life, for treating glucose

XX intolerance associated or not with insulin resistance pathologies, type

XX II diabetes

XX Claim 5; Page 55; 77pp; English.

XX The present invention relates to modified biological peptides with

XX increased biological potency, prolonged activity and/or increased

XX half-life. The peptides of the invention are useful in the treatment

XX of glucose intolerance which may be associated with insulin resistance

XX pathologies, and in the treatment of type II diabetes. They are also

XX useful for treating bone diseases, cancer, diseases related to

CC inflammatory responses, obesity, autism, pervasive developmental  
 CC disorders, hyperproliferative skin diseases, hormone-dependent diseases,  
 CC they can be used for regulating blood glucose, enhancing mucosal  
 CC regeneration in patients with intestinal diseases, inhibition of  
 CC interleukin-8 production, stimulation of acid release, homeostasis,  
 CC regulation of exocrine and endocrine secretions, smooth muscle  
 CC contraction, feeding, blood pressure, body temperature and cell growth,  
 CC regulation of food intake and energy balance, and stimulation of  
 CC pancreatic secretion or cell growth. AAU85971-AAU86019 represent the  
 CC modified biological peptides of the invention.

XX Sequence 49 AA;

Query Match 22.4%; Score 11; DB 23; Length 49;

Best Local Similarity 100.0%; Pred. No. 0.00012;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELDHIGFQ 39  
 DB 29 CDELDHIGFQ 39

RESULT 25

AA017841 standard; Protein; 68 AA.

XX AA017841;

XX 20-AUG-2002 (first entry)

DE Peptide presentation method related vector encoded protein #2.

XX Peptide presentation; host cell surface; antibody isolation;

KM epitope mapping; bacteria; EETI-II.

XX Unidentified.

OS WO200234906-A2.

XX 02-MAY-2002.

XX 25-OCT-2001; 2001WO-DE04009.

XX 26-OCT-2000; 2000DE-1053224.

XX (UYGE-) UNIV GEORG AUGUST GOETTINGEN.

XX Kolmar H, Christmann A, Wentzel A;

XX WPI; 2002-435621/46.

DR N-PSDB; AAL47111.

XX Cell-surface presentation of peptides or proteins, useful e.g. for  
 PT isolating monospecific antibodies, comprises expression of a fusion  
 PT sequence with truncated intimin, in bacteria -

XX Example 1; Fig 2; 44p; German.

XX The present invention relates to a method for presenting peptides or  
 CC proteins on the surface of host bacteria. The method can be used to  
 CC isolate, from polyclonal mixtures, monospecific antibodies that bind  
 CC selectively to a particular surface-bound polypeptide, or more generally  
 CC any specific binding partners, and for epitope mapping. The present  
 CC sequence is a vector encoded protein fragment described in the  
 CC exemplification of the invention.

XX Sequence 68 AA;

Query Match 22.4%; Score 11; DB 23; Length 68;

Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVYPDPL 16

DB 14 LGAPVYPDPL 24

RESULT 26

AAW76097 standard; Protein; 99 AA.

XX AAW76097;

XX 21-DEC-1998 (first entry)

DE Rat osteocalcin protein.

XX Promoter; tissue-specific gene expression; skeletal tissue;

KW stem cell; bone; cartilage; osteocalcin; ROC gene; rat;

KW osteoporosis; osteopenia; osteosarcoma; cancer; metastasis;  
 KW gene therapy.

XX Rattus sp.

XX WO9839427-A2.

XX 11-SEP-1998.

XX 06-MAR-1998; 98WO-US04421.

XX 06-MAR-1997; 97US-0039839.

XX (UYMA-) UNIV MASSACHUSETTS.

XX Frenkel B, Hou Z, Lian JB, Nilsson S, Peters S;

PI Quesenberry P, Stein GS, Stein JL;

XX WPI; 1998-495839/42.

DR N-PSDB; AAV46432.

XX Expression of exogenous genes in differentiated cells - by  
 PT transducing pluripotent stem cells capable of maturing into  
 PT differentiated cells with nucleic acid comprising exogenous gene,  
 PT useful for, e.g. treatment of osteoporosis

XX Disclosure; Page 43-45; 63p; English.

XX This is the amino acid sequence of the rat osteocalcin, deduced  
 CC from the coding exons of the rat ROC gene (see AAV46432). The  
 CC invention pertains to a method for expressing endogenous genes in  
 CC differentiated cells of a specific type. The method involves  
 CC contacting pluripotent stem cells capable of maturing into  
 CC differentiated cells with a nucleic acid comprising an exogenous  
 CC gene linked to a regulatory element capable of controlling expression  
 CC of the exogenous gene in the differentiated cells. A population of  
 CC transduced stem cells capable of maturing into differentiated cells  
 CC expressing the exogenous gene is produced. Preferably, the  
 CC differentiated cells are in a tissue of interest, such as bone or  
 CC cartilage, and the exogenous gene is operably linked to at least one  
 CC osteocalcin regulatory element derived from the human osteocalcin  
 CC gene (hOC) promoter (see AAV46429). The exogenous gene can encode a  
 CC therapeutic protein useful for treating a disease, especially  
 CC osteoporosis, osteopenia, osteosarcoma, primary malignancy or  
 CC metastases (all claimed).

XX Sequence 99 AA;

Query Match 22.4%; Score 11; DB 19; Length 99;

Best Local Similarity 100.0%; Pred. No. 0.00022;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELDHIGFQ 39  
 DB 78 CDELDHIGFQ 88



```

RESULT 27
AAW04605
ID AAW04605 standard; peptide: 13 AA.
XX
AC AAW04605;
XX
DT 13-AUG-1997 (first entry)
XX
DE Osteocalcin 7-19 fragment for mass spectrometry analysis.
XX
KM Mass spectrometry; polymer analysis; biopolymer analysis.
XX
OS Synthetic.
XX
PN W09636986-A1.
XX
PD 21-NOV-1996.
XX
PE 17-MAY-1996; 96WO-US07146.
XX
PR 19-MAY-1995; 95US-0447175.
XX
PR 19-MAY-1995; 95US-0446055.
XX
PA (PERS-) PERSEPTIVE BIOSYSTEMS INC.
XX
PI Patterson DH, Tarr GE;
XX
DR WPI: 1997-012308/01.
XX
PT Sequencing polymers, e.g. DNA, RNA, peptide nucleic acids, proteins,
PT etc. - by obtaining mass to charge ratios of polymer fragments,
PT pref. using mass spectrometer, and performing statistical analysis
XX
PS Example 2; Page 32; 86pp; English.
XX
CC A method of obtaining sequence information about a polymer (e.g. DNA,
CC RNA, peptide nucleic acids, proteins, peptides and carbohydrates)
CC comprising monomers of known mass has been claimed. The present
CC sequence represents a fragment of osteocalcin (7-19), and was used as
CC an example as a digestion before analysis by mass spectrometry,
CC using this novel on-plate strategy. Total sequence information
CC from a nine well digestion can be represented in a single digestion or
CC it is often derived from two or more wells. The methods, apparatus and
CC kit (claimed) can be used for the analysis of polymers, particularly
CC biopolymers, e.g. DNA, RNA, peptide nucleic acids, proteins, peptides
CC and carbohydrates. It provides a rapid, automated and cost effective
CC sequencing of polymers, with a statistical certainty.
CC
SQ Sequence 13 AA:
OY 7 GAPVYPDPL 16
DB 1 GAPVYPDPL 10
Query Match 20.4%; Score 10; DB 18; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

XX
PN US5998170-A.
XX
PD 07-DEC-1999.
XX
PE 03-OCT-1997; 97US-0943915.
XX
PR 03-OCT-1997; 97US-0943915.
XX
PA (AMGE-) AMGEN INC.
XX
PI Arakawa T, Itoh N, Danilenko DM, Martin FH;
XX
DR WPI: 2000-085497/07.
XX
PT Fibroblast growth factor family polypeptide which stimulates
PT proliferation and growth of hepatocytes is useful for treating hepatic
PT disorders -
XX
PS Example III; Column 14; 33pp; English.
XX
CC The invention relates to rat and human fibroblast growth factor-16
CC (FGF-16, AAY58428-Y58429), and nucleotides which encode these proteins.
CC FGF-16 has hepatocyte proliferation and growth activity, and
CC increases hepatic production of triglycerides and serum proteins (e.g.,
CC albumin). FGF-16 nucleic acids and/or proteins may be used for
CC stimulating the proliferation and development of hepatocytes both in
CC vitro and in vivo. The isolated nucleic acid molecules may be used
CC directly in cell or gene therapy applications to treat or prevent liver
CC disorders, including hepatic cirrhosis, fulminant liver failure, damage
CC caused by acute viral hepatitis and toxic insults to the liver.
CC This sequence represents an E tag, DNA encoding which was fused to the
CC 3' end of the rat FGF-16 coding region, along with DNA encoding a
CC hexahistidine tag. The tagged rat FGF-16 cDNA was cloned into a
CC baculovirus expression system in an exemplification of the present
CC invention.
CC
SQ Sequence 13 AA:
OY 7 GAPVYPDPL 16
DB 1 GAPVYPDPL 10
Query Match 20.4%; Score 10; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 28
AAY58430
ID AAY58430 standard; peptide: 13 AA.
XX
AC AAY58430;
XX
DT 27-MAR-2000 (first entry)
XX
DE E tag peptide, SEQ ID NO:6.
XX
KM E tag; fibroblast growth factor; FGF-16; rat;
KM baculovirus expression system.
XX
OS Synthetic.

```

```

RESULT 29
AAG65668
ID AAG65668 standard; peptide: 13 AA.
XX
AC AAG65668;
XX
DT 07-JAN-2002 (first entry)
XX
DE E tag used for recombinant expression of FGF-1like polypeptide.
XX
KM Fibroblast growth factor; FGF; FGF-1like polypeptide; mitogenic;
KM fat deposition; vulnuerary; antiulcer; dermatological; anorectic;
KM antidiabetic; antiinflammatory; cytostatic; hepatic; vitruclide;
KM neuroprotectant; pulmonary; gene therapy; vaccine; human.
XX
OS Synthetic.
XX
PN W0200172957-A2.
XX
PD 04-OCT-2001.
XX
PE 02-APR-2001; 2001WO-IB00664.
XX
PR 31-MAR-2000; 2000US-0540118.

```

PA	(ITOH/)	ITOH N.
XX		
PI	Itch N;	
XX		
DR	WPI: 2001-611623/70.	
XX		
PT	New human nucleic acid encoding fibroblast growth factor-like peptide,	
PT	useful for treatment and diagnosis of e.g. wounds and inflammatory	
PI	bowel disease -	
XX		
ES	Example 3; Page 117, 172pp; English.	
XX		
CC	The invention provides human nucleic acids encoding fibroblast growth	
CC	factor (FGF)-like peptide. The FGF-like polypeptides can be expressed by	
CC	standard recombinant methodology and are mitogenic for a wide range of	
CC	cells, inducing differentiation and proliferation, and inhibiting	
CC	deposition of fat. The FGF-like polypeptides, polynucleotides and	
CC	specific antibodies and modulators are useful for treating a very wide	
CC	range of diseases and conditions, e.g. wounds, ulcers, skin aging,	
CC	hepatitis, diabetes, alopecia, inflammatory bowel disease, emphysema, viral	
CC	hepatitis, multiple sclerosis, respiratory distress syndrome, tumors of	
CC	the eye, etc., also for maintaining organs before transplant and	
CC	supporting culture of primary cells and tissues. Sequences AAG55668-69	
CC	represent amino acid sequences of protein tags used for recombinant	
CC	expression of FGF-like polypeptide.	
XX		
SQ	Sequence 13 AA;	
	Query Match	20.4%; Score 10; DB 22; Length 13;
	Best Local Similarity	100.0%; Pred. No. 0 00047;
	Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	7 GAPVPPDPL 16	
	1 GAPVPPDPL 10	
DB		
	RESULT 30	
	AAU02089	
ID	AAU02089 standard; Peptide; 13 AA.	
XX		
AC	AAU02089;	
XX		
DT	07-SEP-2001 (first entry)	
XX		
DE	E-tag peptide.	
XX		
KW	Fibroblast growth factor 20; FGF-20; Parkinson's disease;	
KW	substantia nigra; dopaminergic neuron; cochlea-associated disease;	
KW	otosclerosis; Cogan's syndrome; Meniere's disease; Pendred's syndrome;	
KW	diabetes-associated hearing loss; congenital malformations;	
KW	autoimmune disease-related hearing loss; age-related hearing loss;	
KW	deafness; ischaemia-related hearing disturbance; immunogen;	
KW	antibody; neuro-degenerative disease; tendonitis; wound healing;	
XX	stroke; ischaemia; E-tag.	
XX		
OS	Synthetic.	
XX		
PN	WO200131008-A2.	
XX		
PD	03-MAY-2001.	
XX		
PF	20-OCT-2000; 2000WO-US29237.	
XX		
PR	22-OCT-1999; 99US-0161162.	
XX		
PR	08-MAR-2000; 2000US-0187856.	
XX		
PA	(CHIR ) CHIRON CORP.	
XX	(KYOU ) UNIV KYOTO.	
XX		
PI	Itch N, Kavanaugh WM;	
XX		
DR	WPI: 2001-308642/32.	

XX Example 1; Page 43; 48pp; English.  
PS  
XX  
CC The present invention relates to a recombinant antibody capable of  
CC specifically binding to sperm agglutination antigen-1 (SAGA-1).  
CC The recombinant antibody has two peptide fragments of the S19  
CC antibody and the fragments are joined together by a linker.  
CC The recombinant monoclonal antibodies are useful in a passive  
CC immunity composition for contraception as they inhibit the ability of  
CC sperm to fertilize an egg. The antibodies may be used e.g. as an  
CC active ingredient of a spermstatic agent, or as a component of a  
CC spermicidal contraceptive.  
XX  
SQ Sequence 13 AA;  
  
Query Match 20.4%; Score 10; DB 22; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00047;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 7 GAVPYPPDPL 16  
Db 1 GAVPYPPDPL 10  
|||||  
|  
  
RESULT 32  
AA015409  
ID AA015409 standard; Peptide: 13 AA.  
XX  
XX AA015409;  
AC  
XX  
XX 27-SEP-2002 (first entry)  
DT  
XX  
DE E tag - peptide tag sequence.  
XX  
XX Mammalian; nuclear receptor cofactor; CF7; CF8; metabolism regulation;  
KW cell homeostasis; cell proliferation; differentiation;  
KW pathological cellular aberration; cellular defence mechanism; E tag.  
XX  
OS Synthetic.  
XX  
XX WO200242322-A2.  
PN  
XX  
PD 30-MAY-2002.  
XX  
XX 21-NOV-2001; 2001WO-EP13548.  
PF  
XX  
XX 21-NOV-2000; 2000EP-0125524.  
PR  
XX  
XX (LION-) LION BIOSCIENCE AG.  
PA  
XX  
XX Jackson D, Casari G, Suckow J;  
PI  
XX  
XX WPI; 2002-566559/60.  
DR  
XX  
XX Novel nuclear receptor cofactors, CF7 and CF8 for identifying  
PT modulators useful for inhibiting cellular function of cofactor and for  
PT treating metabolic disorders, Immunological Indications and hormonal  
PT dysfunction -  
XX  
PS Disclosure; Page 16; 68pp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of two  
CC mammalian nuclear receptor cofactor proteins (CF7 and CF8). The CF7 and  
CC CF8 protein sequences of the invention are useful for screening agents  
CC that are capable of inhibiting the cellular function of cofactor CF7  
CC and/or CF8. The CF7 and CF8 DNA sequences are useful for modulating genes  
CC involved in cellular functions, such as: regulation of metabolism and  
CC cell homeostasis, cell proliferation and differentiation, pathological  
CC cellular aberrations, or cellular defence mechanisms. The present amino  
CC acid sequence represents a peptide tag that was used in the invention.  
XX  
SQ Sequence 13 AA;

Query Match 20.4%; Score 10; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00047;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 7 GAVPYPPDPL 16  
Db 1 GAVPYPPDPL 10  
|||||  
|  
  
RESULT 33  
ABB83453  
ID ABB83453 standard; Peptide: 13 AA.  
XX  
XX ABB83453;  
AC  
XX  
XX 30-SEP-2002 (first entry)  
DT  
XX  
DE E tag.  
XX  
XX Nuclear receptor cofactor; CF11; metabolism regulation;  
KW cell homeostasis; cell proliferation; cell differentiation;  
KW central nervous system; metabolic disorder; immunological disease;  
KW hormonal dysfunction; neuronal disease.  
XX  
OS Synthetic.  
XX  
XX WO200235385-A2.  
PN  
XX  
XX 11-JUL-2002.  
PD  
XX  
XX 21-DEC-2001; 2001WO-EP15133.  
PF  
XX  
XX 30-DEC-2000; 2000EP-0128768.  
PR  
XX  
XX (LION-) LION BIOSCIENCE AG.  
PA  
XX  
XX Jackson D, Casari G, Suckow J;  
PI  
XX  
XX WPI; 2002-583601/62.  
DR  
XX  
XX Novel polypeptide useful for screening agents capable of inhibiting or  
PT activating the cellular function of cofactor of mammalian nuclear  
PT receptor CF11 -  
XX  
PS Disclosure; Page 15; 59pp; English.  
XX  
XX The present invention relates to CF11, a novel human nuclear receptor  
CC cofactor (ABB83451). CF11 modulates genes involved in various cellular  
CC functions such as regulation of metabolism and cell homeostasis, cell  
CC proliferation and differentiation, pathological cellular aberrations or  
CC cellular defence mechanisms. CF11 binds retinoic acid receptor (RAR;  
CC NR1A) and thyroid hormone receptor (TR; NR1B) but not retinoid X receptor  
CC (TR; NR2B) or steroid hormone receptors. CF11 expression is restricted to  
CC the central nervous system and could be confined to neurons in the  
CC dentate gyrus of the hippocampus, the amygdala, thalamic and hypothalamic  
CC regions. Antagonists to CF11 are useful in the development of drugs  
CC against diseases such as metabolic disorders, immunological indications,  
CC hormonal dysfunctions and/or neurosystemic diseases or related to defects  
CC in neuronal diseases. The present sequence is a peptide tag, which may be  
CC used to generate recombinant CF11 proteins.  
XX  
SQ Sequence 13 AA;  
  
Query Match 20.4%; Score 10; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00047;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 7 GAVPYPPDPL 16  
Db 1 GAVPYPPDPL 10  
|||||  
|  
  
RESULT 34

AA098372	standard; Peptide; 13 AA.
AA098372	
AA098372;	
24-SEP-2002	(first entry)
Synthetic E tag	peptide sequence.
Cofactor 9; CF9;	nuclear receptor; metabolic disorder;
drug design;	immunological indication; hormonal dysfunction;
neurosystemic disease;	cofactor binding assay; E tag.
Synthetic.	
WO200244365-A1.	
06-JUN-2002.	
28-NOV-2001;	2001WO-EPI3891.
28-NOV-2000;	2000EP-0126022.
(LION-) LION	BIOSCIENCE AG.
Jackson D, Casari G, Suckow J;	
WPI; 2002-527709/56.	
Novel mammalian nuclear receptor cofactor 9 polypeptide useful for	identifying compounds for treating metabolic disorders, immunological
indications, hormonal dysfunctions and/or neurosystemic diseases	-
Disclosure; Page 16; 62pp; English.	
The present invention relates to a new mammalian nuclear receptor	cofactor 9 (CF9) polypeptide. The invention is useful for screening for
agents which are capable of inhibiting the cellular function of CF9.	
The invention is also useful for screening for nuclear receptors,	
for screening drugs for agonist and antagonist activity and for	
screening drugs useful in regulating physiological responses associated	with CF9, and in structural drug designing. The molecules of the
invention are useful for identifying compounds for treating	metabolic disorders, immunological indications, hormonal dysfunctions
and/or neurosystemic diseases. The invention can be useful as	intermediates for making cellular preparations for cofactor binding
assays, which are useful in drug screening. The present amino acid	sequence represents the E tag peptide sequence, as described in the
invention.	
Sequence 13 AA;	
Query Match	20.4%; Score 10; DB 23; Length 13;
Best Local Similarity	100.0%; Pred. No. 0.00047;
Matches 10; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
7 GAVVPYDPDL 16	
1 GAVVPYDPDL 10	
RESULT 35	
AA078634	
AA078634	standard; Peptide; 13 AA.
AA078634;	
18-JUN-2002	(first entry)
Synthetic E tag	peptide sequence.
Pregnane X; receptor; PXR; cofactor; xenobiotic; E tag.	

OS	Synthetic.
XX	
PN	WO200218420-A2.
PD	
XX	07-MAR-2002.
PF	
XX	17-AUG-2001; 2001WO-EP09488.
PR	
XX	28-AUG-2000; 2000EP-0118634.
PA	(LION-) LION BIOSCIENCE AG.
XX	
PI	Albers M, Ellwanger S, Koejl M, Loeser E;
PT	WPI: 2002-292195/73.
PP	
XX	New nucleic acids and cofactors of the pregnane x nuclear receptor
XX	(PXR), which the nucleic acid encodes, useful for screening agonists or
XX	antagonists of PXR, and for determining a subject's response to
XX	xenobiotic substances or drugs -
PS	
XX	Disclosure: Page 17; 102pp; English.
XX	
CC	The present invention relates to a new nucleic acid molecule and its
CC	encoded polypeptide. The nucleic acid codes for a cofactor of the
CC	pregnane x nuclear receptor (PXR). The polypeptide encoded by the
CC	nucleic acid comprises 225 amino acids or 293 amino acids fully defined
CC	in the specification. The nucleic acid is useful for making vectors and
CC	transforming cells, both of which are ultimately useful for producing the
CC	cofactor proteins. The nucleic acids may also be used for determining a
CC	subject's response to xenobiotic substances or drugs. The proteins or
CC	complexes are useful for screening substances that bind the proteins or
CC	complexes, particularly agonists or antagonists of PXR. The present
CC	amino acid sequence represents the E tag peptide that was used in the
CC	invention to facilitate purification of recombinant proteins. The
CC	cofactor proteins of the invention are useful for screening for PXR.
XX	
XX	
SQ	Sequence 13 AA:
	Query Match 20.4%; Score 10; DB 23; Length 13;
	Best Local Similarity 100.0%; Pred. No. 0.00047;
	Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY	7 GAPVVPDPL 16 
Db	1 GAPVVPDPL 10
RESULT 36	
AAE18828	
ID	AAE18828 standard; peptide; 13 AA.
XX	
AC	AAE18828;
XX	
DT	17-MAY-2002 (first entry)
DE	E tag used in the production of FGF-like polypeptides.
XX	
Fibroblast growth factor; FGF-like protein; wound healing; bullousa;	
KW	epidermolysis; erosive gastritis; inflammatory bowel disease; ulcer;
KW	oesophagitis; Crohn's disease; hyaline membrane disease; emphysema;
KW	pulmonary fibrosis; hepatic cirrhosis; liver failure; angiogenesis;
KW	multiple sclerosis; neurodegenerative disease; lung abnormality;
KW	viral hepatitis; respiratory distress syndrome; tumour; skin aging;
KW	gene therapy; vaccine; human; E tag.
XX	
OS	Homo sapiens.
XX	
PN	US2002001825-A1.
XX	
PD	03-JAN-2002.
XX	
PF	02-APR-2001; 2001US-0822485.

XX 31-MAR-2000; 2000US-0540118.  
XX (ITOH.) ITOH N.  
XX Itoh N;  
XX WPI; 2002-187704/24.  
XX Novel fibroblast growth factor-like polypeptide useful for treating,  
PT ameliorating and/or preventing dermal wounds, gastric ulcer, Crohn's  
PT disease and pulmonary inflammation -  
XX  
XX Example 3; Page 31; 63pp; English.  
XX The invention relates to fibroblast growth factor (FGF)-like  
CC polypeptides and nucleic acid molecules encoding such polypeptides.  
CC Sequences of the invention are useful for treating, preventing or  
CC ameliorating a medical condition. They are useful for treating dermal  
CC wounds, epidermolysis, bullosa, male pattern alopecia, gastric ulcer,  
CC duodenal ulcer, erosive gastritis, oesophagitis, oesophageal reflux  
CC disease, inflammatory bowel disease, Crohn's disease, radiation- or  
CC chemotherapy-induced gut toxicity, hyaline membrane disease, necrosis  
CC of the respiratory epithelium, emphysema, pulmonary inflammation,  
CC pulmonary fibrosis, hepatic cirrhosis, toxic insults to the liver,  
CC fulminant liver failure, viral hepatitis, mucositis, multiple sclerosis  
CC and other neurodegenerative diseases, infantile respiratory distress  
CC syndrome, bronchopulmonary dysplasia, acute respiratory distress  
CC syndrome or other lung abnormalities, tumours of the eye or the other  
CC tissues and organs. FGF-like polypeptides are useful stimulating  
CC angiogenesis, promoting wound healing, modulating differentiation of  
CC neuronal cells, adipocytes and skeletal muscle cells, preventing or  
CC ameliorate skin aging, preventing hair loss, stimulating the growth  
CC and differentiation of haematopoietic cells and bone marrow cells and  
CC maintaining organs before transplantation and for supporting cultures  
CC of primary cells and tissues. Sequences of the invention are also  
CC used in gene therapy and as vaccines. The present sequence is E tag  
CC used in the production of FGF-like polypeptides.  
XX  
SQ Sequence 13 AA;  
Query Match 20.4%; Score 10; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00047;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAPVPRDPL 16  
Db 1 GAPVPRDPL 10  
RESULT 37  
AAM48056  
ID AAM48056 standard; peptide; 13 AA.  
XX  
AC AAM48056;  
XX  
DT 14-MAR-2002 (first entry)  
XX  
DE E-tag sequence.  
XX  
KM Phenotype: protein binding partner; ligand; E-tag.  
XX  
OS Synthetic.  
XX  
PN WO200186297-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 09-MAY-2001; 2001WO-US15092.  
XX  
PR 09-MAY-2000; 2000US-202912P.  
XX  
PA (DGIB-) DGI BIOTECHNOLOGIES INC.

XX Blume AJ, Goldstein N, Pillutla R, Hsiao K, Prendergast J;  
XX WPI; 2002-089808/12.  
XX Identifying a naturally occurring binding partner or binding partner  
PT precursor for a target, comprises comparing identified amino acid  
PT sequence motifs to known sequences of a genome to identify a gene  
PT product of the genome having the motif -  
XX  
XX Example 1; Page 20; 47pp; English.  
XX The invention relates to identifying a naturally occurring binding  
CC partner or binding partner precursor for a target, comprising comparing  
CC the identified amino acid sequence motifs to known amino acid sequences  
CC of a genome and identifying a gene product of the genome possessing the  
CC motif as the naturally occurring binding partner or partner precursor,  
CC for the target. The method is useful for determining the influence which  
CC specific genotypes have on phenotypes and for obtaining peptides which  
CC may be used to identify the natural protein partner of the target and  
CC enable synthesis of peptides which alter the phenotype of cells  
CC expressing the target. The method may also be used to identify peptide  
CC ligands capable of activating or inhibiting gene products through their  
CC ability to bind to such gene products as well as the activity and  
CC function of the gene products themselves. The present sequence is that of  
CC a tag sequence for peptide library sequences useful to the invention.  
XX  
SQ Sequence 13 AA;  
Query Match 20.4%; Score 10; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00047;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAPVPRDPL 16  
Db 1 GAPVPRDPL 10  
RESULT 38  
AAY70693  
ID AAY70693 standard; peptide; 15 AA.  
XX  
AC AAY70693;  
XX  
DT 18-JUL-2000 (first entry)  
XX  
DE E. coli antibody derived peptide to isolate attractin specific scFv.  
XX  
KM Attractin: immune response; macrophage; monocyte; T cell; cancer; scFv;  
KM immunostimulant; immunosuppressed patient; immunodeficiency syndrome;  
KM transplant; autoimmune disease; antibody; single chain variable fragment.  
XX  
OS Escherichia coli.  
XX  
PN WO200015651-A1.  
XX  
PD 23-MAR-2000.  
XX  
PF 14-SEP-1999; 99WO-US20948.  
XX  
PR 14-SEP-1998; 98US-0100137.  
XX  
PA (DAND ) DANA FARBER CANCER INST INC.  
XX  
PI Duke-Cohan JS, Schlossman SF;  
XX  
DR WPI; 2000-271373/23.  
XX  
PT Isolated nucleic acids encoding human attractin polypeptides useful for  
PT enhancing immune responses -  
XX  
PS Example 7; Page 48; 120pp; English.

CC The patent discloses four forms of human attractin polypeptides  
 CC which enhance immune response by promoting macrophage and monocyte  
 CC spreading in the presence of T cells. These include soluble attractin-1  
 CC and -2 and membrane attractin-1 and -2. These various forms of attractin  
 CC are encoded by alternatively spliced mRNA molecule transcribed  
 CC from a single gene. The present sequence is a peptide downstream to  
 CC kappa light chain of an E. coli antibody. Antibody directed against this  
 CC region is used in an affinity column to isolate E. coli soluble single  
 CC chain variable fragment (scFv) specific for attractin. Attractin can be  
 CC used to enhance immune response  
 CC in immunosuppressed patients such as those undergoing chemo- and  
 CC radio-therapy treatment for cancer or those suffering from common  
 CC variable immunodeficiency syndrome. The protein may also be used to  
 CC screen modulators (agonists and antagonists) of immune responses  
 CC which may also be used to regulate immune reactions. Attractin  
 CC antibodies can be used to inhibit immune response in transplant  
 CC recipients or patients afflicted with autoimmune disease.  
 CC  
 SQ Sequence 15 AA;  
 Query Match 20.4%; Score 10; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00054;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 GAPVPPDPL 16  
 Db 1 GAPVPPDPL 10  
 RESULT 39  
 AAY05366  
 ID AAY05366 standard; peptide; 18 AA.  
 XX  
 AC AAY05366;  
 XX  
 DT 30-JUN-1999 (first entry)  
 XX  
 DE HBV specific single stranded antibody fragment.  
 XX  
 KM Single stranded antibody; hepatitis B virus; HBV core protein;  
 KM HBV infection; viral proliferation inhibitor; viral DNA synthesis;  
 KW gene therapy.  
 XX  
 OS Mus sp.  
 OS  
 PN WO9111792-A1.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PF 02-SEP-1998; 98WO-JP03921.  
 XX  
 PR 02-SEP-1997; 97JP-0237054.  
 XX  
 PA (HAYASHI) HAYASHI N.  
 PA (SUMO) SUMITOMO PHARM CO LTD.  
 XX  
 PI Hayashi N, Tohdoh N, Yamamoto H, Yamamoto M;  
 DR WPI: 1999-243623/20.  
 DR N-PSDB; AAX33934.  
 XX  
 XX  
 PT Single-stranded antibody against hepatitis B virus core protein,  
 PT applicable as (gene) therapeutic agents for treatment of hepatitis B  
 PT viral infections  
 XX  
 PS Example 4; Page 68; 72pp; English.  
 XX  
 CC This sequence is the single-stranded antibody of the invention, that  
 CC has the capability of binding to a hepatitis B virus (HBV) core protein.  
 CC Therapeutic agents can be formulated with the antibody for treatment of  
 CC HBV infections by stopping proliferation of the virus through inhibition  
 CC of viral DNA synthesis, and the gene encoding the antibody can be applied  
 CC as an agent to gene therapy.

XX  
 SQ Sequence 18 AA;  
 Query Match 20.4%; Score 10; DB 20; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 0.00063;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 GAPVPPDPL 16  
 Db 4 GAPVPPDPL 13  
 RESULT 40  
 AAE08623  
 ID AAE08623 standard; peptide; 24 AA.  
 XX  
 AC AAE08623;  
 XX  
 DT 01-NOV-2001 (first entry)  
 XX  
 DE Ehtag peptide.  
 XX  
 KW Cytokine; T-cell; interferon-gamma; IFN-gamma; bacterial infection;  
 KW AIDS; diabetic retinopathy; cancer; vaccine; cell-mediated immunity;  
 KW p35; p40; Ehtag.  
 XX  
 OS Synthetic.  
 OS  
 PN WO200140257-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 30-NOV-2000; 2000WO-US32664.  
 XX  
 PR 02-DEC-1999; 99US-0169035.  
 PR 28-NOV-2000; 2000US-0169035.  
 XX  
 PA (MAXY-) MAXYGEN INC.  
 XX  
 PI Leong SR, Punnonen J;  
 DR WPI: 2001-502381/55.  
 XX  
 XX  
 PT Nucleic acids encoding modified cytokine polypeptides (i.e. modified  
 PT p40 and p35 polypeptides), useful for treating various diseases e.g.  
 PT diabetic retinopathy and cancer -  
 XX  
 PS Example 1; Page 143; 223pp; English.  
 XX  
 CC The invention relates to nucleic acids encoding modified cytokine  
 CC polypeptides (i.e. modified p40 and p35 polypeptides). The modified  
 CC cytokine polypeptides have T-cell proliferative and interferon (IFN)-  
 CC gamma induction activities. The polypeptides and polynucleotides of the  
 CC invention are useful for treating various diseases e.g. bacterial  
 CC infections, AIDS, diabetic retinopathy and cancer. The polypeptide and  
 CC polynucleotide may also be useful as a vaccine adjuvant, to enhance a  
 CC vaccinated host's cell-mediated immunity for protective response to a  
 CC pathogen. The present sequence is Ehtag peptide used for  
 CC purification of human wild-type or modified p35 polypeptides.  
 CC  
 SQ Sequence 24 AA;  
 Query Match 20.4%; Score 10; DB 22; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 0.0008;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 GAPVPPDPL 16  
 Db 4 GAPVPPDPL 13  
 RESULT 41  
 AA088451

```

ID  AAU88451 standard; peptide; 38 AA.
XX
AC  AAU88451;
XX
DT  18-JUN-2002 (first entry).
XX
DE  Insulin/insulin-like growth factor receptor-binding peptide #407.
XX
KW  Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
KW  ophthalmological; insulin; receptor; gene therapy; diabetes;
KW  insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
KW  diabetic retinopathy; neurological diseases; stroke;
KW  diabetic neuropathy.
XX
OS  Synthetic.
XX
PN  WO200172771-A2.
XX
PD  04-OCT-2001.
XX
PF  29-MAR-2000; 2000WO-US08528.
XX
PR  29-MAR-2000; 2000WO-US08528.
XX
PA  (DGIB-) DGI BIOTECHNOLOGIES LLC.
PA  (NOVO ) NOVO NORDISK AS.
XX
PI  Beasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;
PI  Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;
PI  Hansen PH, Ravera M, Hsiao K;
XX
DR  WPI; 2002-025774/03.
XX
PT  Modulating insulin activity in mammalian cells, for treating e.g.
PT  diabetes and tumours, comprises using peptides that bind to insulin or
PT  insulin-like growth factor receptors -
XX
PS  Example 28; Page 132; 390pp; English.
XX
CC  The invention relates to a method of modulating insulin activity in
CC  mammalian cells by administering a peptide that binds the insulin
CC  receptor (IR). A composition containing a peptide, optionally expressed
CC  from gene therapy vectors, that binds to site 1 of IR and an insulin
CC  agonist are useful for treating diabetes. Also, peptides that are
CC  antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
CC  useful for treating insulin-like growth factor (IGF)-sensitive tumours
CC  (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
CC  receptor agonists are useful for treating neurological diseases,
CC  including stroke and diabetic neuropathy. The peptides are also useful in
CC  screening for compounds that bind to IR or IGF-1 receptor, potential
CC  therapeutics and research reagents. AAU88034-AAU90957 represent IR
CC  and/or IGF-1 receptor-binding peptides and related amino acid sequences
CC  of the invention.
XX
SQ  Sequence 38 AA;
XX
Query Match 20.4%; Score 10; DB 23; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 GAPVPPDPL 16
DB 24 GAPVPPDPL 33

```

```

DE  Insulin/insulin-like growth factor receptor-binding peptide #398.
XX
KW  Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
KW  ophthalmological; insulin; receptor; gene therapy; diabetes;
KW  insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
KW  diabetic retinopathy; neurological diseases; stroke;
KW  diabetic neuropathy.
XX
OS  Synthetic.
XX
PN  WO200172771-A2.
XX
PD  04-OCT-2001.
XX
PF  29-MAR-2000; 2000WO-US08528.
XX
PR  29-MAR-2000; 2000WO-US08528.
XX
PA  (DGIB-) DGI BIOTECHNOLOGIES LLC.
PA  (NOVO ) NOVO NORDISK AS.
XX
PI  Beasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;
PI  Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;
PI  Hansen PH, Ravera M, Hsiao K;
XX
DR  WPI; 2002-025774/03.
XX
PT  Modulating insulin activity in mammalian cells, for treating e.g.
PT  diabetes and tumours, comprises using peptides that bind to insulin or
PT  insulin-like growth factor receptors -
XX
PS  Example 28; Page 131; 390pp; English.
XX
CC  The invention relates to a method of modulating insulin activity in
CC  mammalian cells by administering a peptide that binds the insulin
CC  receptor (IR). A composition containing a peptide, optionally expressed
CC  from gene therapy vectors, that binds to site 1 of IR and an insulin
CC  agonist are useful for treating diabetes. Also, peptides that are
CC  antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
CC  useful for treating insulin-like growth factor (IGF)-sensitive tumours
CC  (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
CC  receptor agonists are useful for treating neurological diseases,
CC  including stroke and diabetic neuropathy. The peptides are also useful in
CC  screening for compounds that bind to IR or IGF-1 receptor, potential
CC  therapeutics and research reagents. AAU88034-AAU90957 represent IR
CC  and/or IGF-1 receptor-binding peptides and related amino acid sequences
CC  of the invention.
XX
SQ  Sequence 39 AA;
XX
Query Match 20.4%; Score 10; DB 23; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 GAPVPPDPL 16
DB 25 GAPVPPDPL 34

```

```

RESULT 42
AAU88442
ID  AAU88442 standard; peptide; 39 AA.
XX
AC  AAU88442;
XX
DT  18-JUN-2002 (first entry)
XX

```

```

RESULT 43
AAU19897
ID  AAU19897 standard; protein; 43 AA.
XX
AC  AAU19897;
XX
DT  07-DEC-1997 (first entry)
XX
DE  Vector pUC119EHIS-encoded polypeptide.
XX
KW  Carcinoembryonic antigen; CEA; human; antibody; scfv;
KW  tumour marker; lung cancer; breast cancer; colon cancer;
KW  adenocarcinoma; diagnosis; vector; plasmid pUC119EHIS.
XX

```

```

OS Synthetic.
XX
XX Key Location/Qualifiers
XX 23..34
XX Peptide /label= E_tag
XX 38..43
XX Peptide /label= His_tag
XX
XX WO9720932-A1.
XX
XX 12-JUN-1997.
XX
XX 09-DEC-1996; 96WO-GB03043.
XX
XX 11-OCT-1996; 96GB-0021295.
XX 07-DEC-1995; 95GB-0025004.
XX 23-MAY-1996; 96GB-0010824.
XX
XX (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX Allen DJ, McCafferty JG, Osbourn JK.
XX
XX WPI: 1997-319779/29.
XX N-PSDB: AAT72137.
XX
XX Specific binding members for human carcinoembryonic antigen - bind
XX to the A3-83 extracellular domain of hCEA and are substantially
XX non-cross-reactive with human liver cells; used for diagnosing
XX cancer
XX
XX Example 7; Fig 11; 128pp; English.
XX
XX This polypeptide sequence is encoded by the cloning site region
XX (see AAT72137) in vector pUC119EHTS. It includes an E tag for
XX detection with anti-E tag antibodies and a hexahistidine tag for
XX IMAC purification. Human carcinoembryonic antigen (CEA) epitope
XX sequences were cloned into the vector for preparation of purified
XX CEA domains. These were used to examine the domain recognition of
XX anti-CEA antibody CEA6 (see AAW19881) and To6D11.
XX
XX Sequence 43 AA:
SQ
Query Match 20.4%; Score 10; DB 18; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 GAVPYPPDPL 16
DB 23 GAVPYPPDPL 32

```

```

PF 29-MAR-2000; 2000WO-US08528.
XX
XX 29-MAR-2000; 2000WO-US08528.
XX
XX (DGIB-) DGI BIOTECHNOLOGIES LLC.
XX (NOVO ) NOVO NORDISK AS.
XX
XX Beasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;
XX Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;
XX Hansen PH, Ravera M, Hsiao K;
XX WPI: 2002-025774/03.
XX
XX Modulating insulin activity in mammalian cells, for treating e.g.
XX diabetes and tumours, comprises using peptides that bind to insulin or
XX insulin-like growth factor receptors
XX
XX Example 28; Page 131; 390pp; English.
XX
XX The invention relates to a method of modulating insulin activity in
XX mammalian cells by administering a peptide that binds the insulin
XX receptor (IR). A composition containing a peptide, optionally expressed
XX from gene therapy vectors, that binds to site 1 of IR and an insulin
XX agonist are useful for treating diabetes. Also, peptides that are
XX antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
XX useful for treating insulin-like growth factor (IGF)-sensitive tumours
XX (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
XX receptor agonists are useful for treating neurological diseases.
XX including stroke and diabetic neuropathy. The peptides are also useful in
XX screening for compounds that bind to IR or IGF-1 receptor, potential
XX therapeutics and research reagents. AAU88034-AAU90957 represent IR
XX and/or IGF-1 receptor-binding peptides and related amino acid sequences
XX of the invention.
XX
XX Sequence 44 AA:
SQ
Query Match 20.4%; Score 10; DB 23; Length 44;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 GAVPYPPDPL 16
DB 30 GAVPYPPDPL 39

```

```

RESULT 44
AAU88438
ID AAU88438 standard; Peptide; 44 AA.
XX
XX AAU88438;
XX
XX 18-JUN-2002 (first entry)
XX
XX Insulin/insulin-like growth factor receptor-binding peptide #394.
XX
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
XX ophthalmological; insulin; receptor; gene therapy; diabetes;
XX insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
XX diabetic retinopathy; neurological diseases; stroke;
XX diabetic neuropathy.
XX
XX Synthetic.
XX
XX WO200172771-A2.
XX
XX 04-OCT-2001.
XX
XX
XX

```

```

RESULT 45
AAU88447
ID AAU88447 standard; Peptide; 44 AA.
XX
XX AAU88447;
XX
XX 18-JUN-2002 (first entry)
XX
XX Insulin/insulin-like growth factor receptor-binding peptide #403.
XX
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
XX ophthalmological; insulin; receptor; gene therapy; diabetes;
XX insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
XX diabetic retinopathy; neurological diseases; stroke;
XX diabetic neuropathy.
XX
XX Synthetic.
XX
XX WO200172771-A2.
XX
XX 04-OCT-2001.
XX
XX 29-MAR-2000; 2000WO-US08528.
XX
XX 29-MAR-2000; 2000WO-US08528.
XX
XX (DGIB-) DGI BIOTECHNOLOGIES LLC.
XX (NOVO ) NOVO NORDISK AS.
XX
XX
XX

```



XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;  
 PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;  
 PI Hansen PH, Ravera M, Hsiao K;  
 XX  
 DR WPI; 2002-025774/03.

XX  
 PT Modulating insulin activity in mammalian cells, for treating e.g.  
 PT diabetes and tumours, comprises using peptides that bind to insulin or  
 PT insulin-like growth factor receptors  
 XX  
 PS  
 XX

Example 28: Page 131; 390pp: English.

CC The invention relates to a method of modulating insulin activity in  
 CC mammalian cells by administering a peptide that binds the insulin  
 CC receptor (IR). A composition containing a peptide, optionally expressed  
 CC from gene therapy vectors, that binds to site 1 of IR and an insulin  
 CC agonist are useful for treating diabetes. Also, peptides that are  
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
 CC receptor agonists are useful for treating neurological diseases,  
 CC including stroke and diabetic neuropathy. The peptides are also useful in  
 CC screening for compounds that bind to IR or IGF-1 receptor, potential  
 CC therapeutics and research reagents. AU88034-AU90957 represent IR  
 CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
 CC of the invention.  
 XX  
 XX

SQ Sequence 44 AA;

Query Match 20.4%; Score 10; DB 23; Length 44;

Best Local Similarity 100.0%; Pred. No. 0.0013;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVPYPDPL 16  
 |||||  
 Db 30 GAPVPYPDPL 39

Search completed: December 4, 2002, 15:49:41  
 Job time : 36 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:39:05 ; Search time 13 Seconds

(without alignments)  
230.896 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 49  
Sequence: 1 ILYQWLGAPVPPDLEPRR.....DELADHIGFOEAYRRFGPV 49

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 193982 seqs, 61258239 residues

Word size : 0

Total number of hits satisfying chosen parameters: 193982

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Pending Patents AA\_New.\*

1: /cgn2\_6/ptodata/2/paa/US06\_NEM\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US07\_NEM\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US08\_NEM\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US09\_NEM\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US10\_NEM\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEM\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEM\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	100	US-10-283-656-1	Sequence 1, Appl 1
2	42	85.7	58	US-09-724-676-90318	Sequence 90318, A
3	42	85.7	58	US-09-724-676A-90318	Sequence 90318, A
4	13	28.5	13	US-09-802-154-21	Sequence 21, Appl 1
5	13	26.5	257	US-10-096-246-2	Sequence 2, Appl 1
6	7	14.3	433	US-10-278-945-2	Sequence 6123, Ap
7	6	12.2	56	PCR-US02-32727-6123	Sequence 30030, A
8	6	12.2	68	PCR-US02-32727-30030	Sequence 4591, Ap
9	6	12.2	136	PCR-US02-32727-4591	Sequence 17243, A
10	6	12.2	161	PCR-US02-32727-17243	Sequence 14821, A
11	6	12.2	220	PCR-US02-32727-14821	Sequence 13, Appl 1
12	6	12.2	451	US-10-258-662-13	Sequence 8047, Ap
13	6	12.2	474	PCR-US02-32727-8047	Sequence 29930, A
14	6	12.2	509	PCR-US02-32727-29930	Sequence 21198, A
15	6	12.2	515	PCR-US02-32727-21198	Sequence 63352, A
16	6	12.2	600	PCR-US02-32727-25938	Sequence 63352, A
17	6	12.2	694	US-09-724-676-63352	Sequence 63352, A
18	6	12.2	694	US-09-724-676-63353	Sequence 63352, A
19	6	12.2	694	US-09-724-676A-63352	Sequence 63352, A
20	6	12.2	694	US-09-724-676A-63353	Sequence 63356, A
21	6	12.2	902	US-09-724-676-63356	Sequence 63356, A
22	6	12.2	902	US-09-724-676A-63356	Sequence 31, Appl 1
23	6	12.2	934	PCR-US02-32637-31	Sequence 31, Appl 1
24	6	12.2	934	US-10-270-839-31	Sequence 63354, A
25	6	12.2	957	US-09-724-676-63354	Sequence 63354, A
26	6	12.2	957	US-09-724-676A-63354	Sequence 63354, A

27	6	12.2	958	US-09-724-676-87329	Sequence 87329, A
28	6	12.2	958	US-09-724-676A-87329	Sequence 87329, A
29	6	12.2	1021	US-09-724-676-87327	Sequence 87327, A
30	6	12.2	1021	US-09-724-676A-87327	Sequence 87327, A
31	6	12.2	1183	US-09-724-676-87328	Sequence 87328, A
32	6	12.2	1183	US-09-724-676A-87328	Sequence 87328, A
33	6	12.2	1484	US-09-724-676-79242	Sequence 79242, A
34	6	12.2	1484	US-09-724-676A-79242	Sequence 79242, A
35	6	12.2	1484	US-09-724-676A-79242	Sequence 79242, A
36	6	12.2	1493	US-09-724-676A-79252	Sequence 79252, A
37	6	12.2	1493	US-09-724-676-79252	Sequence 79252, A
38	6	12.2	1493	US-09-724-676A-79252	Sequence 79252, A
39	6	12.2	1493	US-09-724-676A-79245	Sequence 79245, A
40	6	12.2	1493	US-09-724-676A-79256	Sequence 79256, A
41	6	12.2	1836	US-09-724-676-79241	Sequence 79241, A
42	6	12.2	1836	US-09-724-676-79251	Sequence 79251, A
43	6	12.2	1836	US-09-724-676A-79241	Sequence 79241, A
44	6	12.2	1836	US-09-724-676A-79251	Sequence 79251, A
45	6	12.2	1845	US-09-724-676-79244	Sequence 79244, A

## ALIGNMENTS

RESULT 1  
US-10-283-656-1  
Sequence 1, Application US/10283656  
GENERAL INFORMATION:  
APPLICANT: EKEMA, George Mbella  
APPLICANT: MAYS, Robert W.  
APPLICANT: BRUDEN, Kurt R.  
TITLE OF INVENTION: Methods for Using Osteocalcin  
FILE REFERENCE: ATX-005  
CURRENT APPLICATION NUMBER: US/10/283,656  
CURRENT FILING DATE: 2002-10-29  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-283-656-1

Query Match 100.0%, Score 49; DB 6; Length 100;  
Best Local Similarity 100.0%, Pred. No. 5e-45;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILYQWLGAPVPPDLEPRREVCENPDCDELADHIGFOEAYRRFGPV 49  
DB 52 ILYQWLGAPVPPDLEPRREVCENPDCDELADHIGFOEAYRRFGPV 100

RESULT 2  
US-09-724-676-90318  
Sequence 90318, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: CompuGen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 CompuGen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 90318  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676-90318

Query Match 85.7%, Score 42; DB 5; Length 58;  
Best Local Similarity 100.0%, Pred. No. 8.5e-38;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APVPYDPLERREVCELNPDCELADHIGFQEA YRRFGPV 49  
 |||||  
 DB 17 APVPYDPLERREVCELNPDCELADHIGFQEA YRRFGPV 58

```

RESULT 3
US-09-724-676A-90318
; Sequence 90318, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90318
;
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-90318

```

```

Query Match      85.7%; Score 42; DB 5; Length 58
Best Local Similarity 100.0%; Pred. No. 8.5e-38;
Matches 42; Conservative 0; Mismatches 0; Incls

QY      8 AAVPYPPDLPERRRCELNPDDDELADHTRGDEARRRFGPV 49
        |||||
Db      17 AAVPYPPDLPERRRCELNPDDDELADHTRGDEARRRFGPV 58

```

```

RESULT 4
US-09-802-154-21
: Sequence 21, Application US/09802154
: GENERAL INFORMATION:
: APPLICANT: Itoh, Nobuyuki
: APPLICANT: Kanauegh, W. Michael
: TITLE OF INVENTION: HUMAN FGF-23 GENE AND GENE EXPRESSION
: TITLE OF INVENTION: PRODUCTS
: FILE REFERENCE: JP-17149, 001/201130, 409
: CURRENT APPLICATION NUMBER: US/09/802,154
: CURRENT FILING DATE: 2001-03-07
: NUMBER OF SEQ ID NOS: 46
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 21
: LENGTH: 13
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: E tag
US-09-802-154-21

```

Query Match Similarity	26.5%	Score 13	DB 5	Length 13
Best Local Similarity	100.0%	Pred. No.	1	4e-07
Matches 13; Conservative	0;	Mismatches	0;	Indels 0;
QY	7	GAPVPPDPLEPR	19	
DB	1	GAPVPPDPLEPR	13	

```

: RESULT 5
: US-10-096-246-2
: Sequence 2, Application US/10096246
: GENERAL INFORMATION:
: APPLICANT: The Minister of National Defence, Government of Canada
: APPLICANT: Fulton, R E
: APPLICANT: Alvi, Azhar E
: APPLICANT: Nagata, Leslie
: TITLE OF INVENTION: Cloning, Expression, Sequencing, and Functional Enhancement of M
: TITLE OF INVENTION: scFv Antibody Against Venezuelan Equine Encephalitis Virus (Vee)
: FILE REFERENCE: NEI-0007
: CURRENT APPLICATION NUMBER: US/10/096,246

```

```

? CURRENT FILING DATE: 2002-03-13
? NUMBER OF SEQ ID NOS: 37
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO: 2
? LENGTH: 257
? TYPE: PRT
? ORGANISM: Mouse hybridoma cell line 1A4A1
US-10-096-246-2

```

Query Match	26.5%;	Score 13;	DB 6;	Length 257;
Best Local Similarity	100.0%;	Pred. No.	2e-06;	
Matches 13;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

```

RESULT 6 US-10-278-945-2
: Sequence 2, Application US/10278945
: GENERAL INFORMATION:
: APPLICANT: Cao et al.
: TITLE OF INVENTION: Retinoic Acid Receptor Epsilon
: FILE REFERENCE: PFI23D1C1
: CURRENT APPLICATION NUMBER: US/10/278,945
: CURRENT FILING DATE: 2002-10-24
: PRIOR APPLICATION NUMBER: US 09/022,789
: PRIOR FILING DATE: 1998-02-12
: PRIOR APPLICATION NUMBER: US 08/466,120
: PRIOR FILING DATE: 1995-06-06
: PRIOR APPLICATION NUMBER: PCT/US94/07266
: PRIOR FILING DATE: 1994-06-24
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 2
: LENGTH: 433
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-278-945-2

```

Query Match Similarity	14.3%	Score 7;	DB 6;	Length 433;
Best Local Similarity	100.0%	Pred. No. 7.6;		
Matches	7;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
0Y	5 WTGAPVP	11		
Db	4 WTGAPVP	10		

```

/ RESULT 7
/ PCT-US02-32727-6123
/ : Sequence 6123, Application PC/TUS0232727
/ : GENERAL INFORMATION:
/ : APPLICANT: Mitcham, Jennifer
/ : APPLICANT: Skeiky, Yasir
/ : APPLICANT: Persing, David
/ : APPLICANT: Bhatia, Ajay
/ : APPLICANT: Maisonneuve, Jean Francois
/ : APPLICANT: Zhang, Yanni
/ : APPLICANT: Wang, Siding
/ : APPLICANT: Jen, Shyian
/ : APPLICANT: Lodes, Michael
/ : APPLICANT: Benson, Darin
/ : APPLICANT: Jones, Robert
/ : APPLICANT: Carter, Darlick
/ : APPLICANT: Barth, Brenda
/ : APPLICANT: Douglaas, John
/ : TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
/ : FILE REFERENCE: 210121.514C1
/ : CURRENT APPLICATION NUMBER: PCT/US02/32727
/ : CURRENT FILING DATE: 2002-10-11
/ : NUMBER OF SEQ ID NOS: 30992

```

```
; SEQ ID NO 6123
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-6123
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 15 PLEPRR 20
      |||||
Db 49 PLEPRR 54
```

```
RESULT 8
PCT-US02-32727-30030
; Sequence 30030, Application PC/TUS0232727
; GENERAL INFORMATION:
```

```
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yannl
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglas, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514c1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 30030
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-30030
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 68;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 15 PLEPRR 20
      |||||
Db 61 PLEPRR 66
```

```
RESULT 9
PCT-US02-32727-4591
; Sequence 4591, Application PC/TUS0232727
; GENERAL INFORMATION:
```

```
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yannl
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglas, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514c1
```

```
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 4591
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-4591
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 136;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 18 PREVC 23
      |||||
Db 92 PREVC 97
```

```
RESULT 10
PCT-US02-32727-17243
; Sequence 17243, Application PC/TUS0232727
; GENERAL INFORMATION:
```

```
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yannl
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglas, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514c1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 17243
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-17243
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 13 PDPLEP 18
      |||||
Db 150 PDPLEP 155
```

```
RESULT 11
PCT-US02-32727-14821
; Sequence 14821, Application PC/TUS0232727
; GENERAL INFORMATION:
```

```
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yannl
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
```

```
; APPLICANT: Douglass,John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 14821
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-14821

Query Match          12.2%; Score 6; DB 1; Length 220;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
DB 166 LGAPVP 171

RESULT 12
US-10-258-662-13
; Sequence 13, Application US/10258662
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAU, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: LU, Dyrung Alina M.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YAO, Monique G.
; APPLICANT: BURFORD, Neil
; APPLICANT: BATRA, Sajeev
; APPLICANT: POLICKY, Jennifer J.
; TITLE OF INVENTION: RNA METABOLISM PROTEINS
; FILE REFERENCE: PF-0771 USN
; CURRENT APPLICATION NUMBER: US/10/258,662
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/201,875
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/200,184
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,090
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/210,232
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 60/220,553
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 2472784CD1
US-10-258-662-13

Query Match          12.2%; Score 6; DB 6; Length 451;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
DB 49 LGAPVP 54

RESULT 13
```

```
PCT-US02-32727-8047
; Sequence 8047, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siging
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darriek
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass,John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 8047
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-8047

Query Match          12.2%; Score 6; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYRR 44
DB 365 QEAYRR 370

RESULT 14
PCT-US02-32727-29930
; Sequence 29930, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siging
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darriek
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass,John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 29930
; SEQ ID NO 29930
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-29930

Query Match          12.2%; Score 6; DB 1; Length 509;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYRR 44
DB 400 QEAYRR 405
```

RESULT 15  
PCT-US02-32727-21198  
; Sequence 21198, Application PC/TUS0232727  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yann  
; APPLICANT: Wang, Siqiang  
; APPLICANT: Jen, Shyian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darlick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglass, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; CURRENT FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 21198  
; LENGTH: 515  
; TYPE: PRT  
; ORGANISM: Propionl acnes  
PCT-US02-32727-21198

Query Match 12.2%; Score 6; DB 1; Length 515;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 APVPYP 13  
Db 210 APVPYP 215

RESULT 16  
PCT-US02-32727-25938  
; Sequence 25938, Application PC/TUS0232727  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yann  
; APPLICANT: Wang, Siqiang  
; APPLICANT: Jen, Shyian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darlick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglass, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; CURRENT FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 25938  
; LENGTH: 600  
; TYPE: PRT  
; ORGANISM: Propionl acnes  
PCT-US02-32727-25938

Query Match 12.2%; Score 6; DB 1; Length 600;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 DCEELA 33  
Db 32 DCEELA 37

RESULT 17  
US-09-724-676-63352  
; Sequence 63352, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn Version 3.2  
; SEQ ID NO 63352  
; LENGTH: 694  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-63352

Query Match 12.2%; Score 6; DB 5; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
Db 19 LGAPVP 24

RESULT 18  
US-09-724-676-63353  
; Sequence 63353, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn Version 3.2  
; SEQ ID NO 63353  
; LENGTH: 694  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-63353

Query Match 12.2%; Score 6; DB 5; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
Db 19 LGAPVP 24

RESULT 19  
US-09-724-676A-63352  
; Sequence 63352, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn Version 3.2  
; SEQ ID NO 63352  
; LENGTH: 694  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-63352

Query Match 12.2%; Score 6; DB 5; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
|||||  
Db 19 LGAPVP 24

RESULT 20  
US-09-724-676A-63353

; Sequence 63353, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 63353  
; LENGTH: 694  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-724-676A-63353

Query Match 12.2%; Score 6; DB 5; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
|||||  
Db 19 LGAPVP 24

RESULT 21

US-09-724-676-63356  
; Sequence 63356, Application US/09724676

; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 63356  
; LENGTH: 902  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-724-676-63356

Query Match 12.2%; Score 6; DB 5; Length 902;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
|||||  
Db 227 LGAPVP 232

RESULT 22  
US-09-724-676A-63356

; Sequence 63356, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 63356  
; LENGTH: 902  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-63356

Query Match 12.2%; Score 6; DB 5; Length 902;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
|||||  
Db 227 LGAPVP 232

RESULT 23

PCT-US02-32637-31  
; Sequence 31, Application PC/TUS0232637

; GENERAL INFORMATION:  
; APPLICANT: Morphotek, Inc.  
; TITLE OF INVENTION: Genetic Hypermutability of Plants for Gene Discovery and Diag  
; FILE REFERENCE: AG0002PCT (MOR-0132)  
; CURRENT APPLICATION NUMBER: PCT/US02/32637  
; CURRENT FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/328,750  
; PRIOR FILING DATE: 2001-10-12  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 934  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-32637-31

Query Match 12.2%; Score 6; DB 1; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12  
|||||  
Db 614 GAPVPY 619

RESULT 24

US-10-270-839-31  
; Sequence 31, Application US/10270839

; GENERAL INFORMATION:  
; APPLICANT: Chao, Qimin  
; APPLICANT: Grasso, Luigi  
; APPLICANT: Sass, Philip M.  
; APPLICANT: Nicolaides, Nicholas C.  
; TITLE OF INVENTION: Genetic Hypermutability of Plants for Gene Discovery and Diag  
; FILE REFERENCE: AG0002US (MOR-0133)  
; CURRENT APPLICATION NUMBER: US/10/270,839  
; CURRENT FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/328,750  
; PRIOR FILING DATE: 2001-10-12  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 934  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-270-839-31

Query Match 12.2%; Score 6; DB 6; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12  
|||||  
Db 614 GAPVPY 619



```
RESULT 25
US-09-724-676-63354
; Sequence 63354, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63354
; LENGTH: 957
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-63354

Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 957;
Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
Db 282 LGAPVP 287

RESULT 26
US-09-724-676A-63354
; Sequence 63354, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63354
; LENGTH: 957
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-63354

Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 957;
Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
Db 282 LGAPVP 287

RESULT 27
US-09-724-676-87329
; Sequence 87329, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87329
; LENGTH: 958
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-87329

Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 958;
Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 LGAPVP 11
Db 94 LGAPVP 99

RESULT 28
US-09-724-676A-87329
; Sequence 87329, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87329
; LENGTH: 958
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-87329

Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 958;
Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
Db 94 LGAPVP 99

RESULT 29
US-09-724-676-87327
; Sequence 87327, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87327
; LENGTH: 1021
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-87327

Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 1021;
Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
Db 94 LGAPVP 99

RESULT 30
US-09-724-676A-87327
; Sequence 87327, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87327
; LENGTH: 1021
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-87327

Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 1021;
Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

US-09-724-676A-87327

Query Match 12.2%; Score 6; DB 5; Length 1021;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11

|||||

DB 94 LGAPVP 99

RESULT 31  
US-09-724-676-87328

; Sequence 87328, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO: 87328

; LENGTH: 1183

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-87328

Query Match 12.2%; Score 6; DB 5; Length 1183;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11

|||||

DB 94 LGAPVP 99

RESULT 32  
US-09-724-676A-87328

; Sequence 87328, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO: 87328

; LENGTH: 1183

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-87328

Query Match 12.2%; Score 6; DB 5; Length 1183;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11

|||||

DB 94 LGAPVP 99

RESULT 33  
US-09-724-676-79242

; Sequence 79242, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO: 79242

; LENGTH: 1484

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-79242

Query Match

Best Local Similarity 100.0%; Score 6; DB 5; Length 1484;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36

|||||

DB 1332 ELADHI 1337

RESULT 34  
US-09-724-676-79252

; Sequence 79252, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO: 79252

; LENGTH: 1484

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-79252

Query Match 12.2%; Score 6; DB 5; Length 1484;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36

|||||

DB 1332 ELADHI 1337

RESULT 35  
US-09-724-676A-79242

; Sequence 79242, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO: 79242

; LENGTH: 1484

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-79242

Query Match

Best Local Similarity 100.0%; Score 6; DB 5; Length 1484;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36

|||||

DB 1332 ELADHI 1337

RESULT 36  
US-09-724-676A-79252

; Sequence 79252, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

```
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; SEQUENCE 79245, Application US/09724676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79252
; LENGTH: 1484
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-79252
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 1484;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 31 ELADHI 36
Db 1341 ELADHI 1337
```

```
RESULT 37
US-09-724-676-79245
; SEQUENCE 79245, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79245
; LENGTH: 1493
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-79245
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 1493;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 31 ELADHI 36
Db 1341 ELADHI 1346
```

```
RESULT 38
US-09-724-676-79256
; SEQUENCE 79256, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79256
; LENGTH: 1493
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-79256
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 1493;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 31 ELADHI 36
Db 1341 ELADHI 1346
```

```
RESULT 39
US-09-724-676A-79245
; SEQUENCE 79245, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79245
; LENGTH: 1493
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-79245
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 1493;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 31 ELADHI 36
Db 1341 ELADHI 1346
```

```
RESULT 40
US-09-724-676A-79256
; SEQUENCE 79256, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79256
; LENGTH: 1493
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676A-79256
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 1493;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 31 ELADHI 36
Db 1341 ELADHI 1346
```

```
RESULT 41
US-09-724-676-79241
; SEQUENCE 79241, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79241
; LENGTH: 1836
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-724-676-79241
```

```
Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 1836;
Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 31 ELADHI 36  
 Db 1332 ELADHI 1337

# RESULT 42

US-09-724-676-79251  
 ; Sequence 79251, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 79251  
 ; LENGTH: 1836  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-79251

Query Match 12.2%; Score 6; DB 5; Length 1836;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
 Db 1332 ELADHI 1337

# RESULT 43

US-09-724-676A-79241  
 ; Sequence 79241, Application US/09724676A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676A  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 79241  
 ; LENGTH: 1836  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676A-79241

Query Match 12.2%; Score 6; DB 5; Length 1836;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
 Db 1332 ELADHI 1337

# RESULT 44

US-09-724-676A-79251  
 ; Sequence 79251, Application US/09724676A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676A  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 79251  
 ; LENGTH: 1836  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676A-79251

Query Match 12.2%; Score 6; DB 5; Length 1836;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
 Db 1332 ELADHI 1337

# RESULT 45

US-09-724-676-79244  
 ; Sequence 79244, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 79244  
 ; LENGTH: 1845  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-79244

Query Match 12.2%; Score 6; DB 5; Length 1845;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
 Db 1341 ELADHI 1346

Search completed: December 4, 2002, 15:43:15  
 Job time : 15 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:37:30 ; Search time 29 Seconds  
(without alignments)  
348.149 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 49  
Sequence: 1 YLYQWLGAIPYPPDPLEPRR.....DELADHIGFEAVRRFYGPV 49

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.podent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	26.5	34	6	09TT28
2	13	26.5	147	11	0925S3
3	13	26.5	170	11	0925S2
4	13	26.5	298	11	09QYFO
5	8	16.3	205	10	09XK33
6	8	16.3	536	12	039303
7	7	14.3	97	13	090VW2
8	7	14.3	268	2	069068
9	7	14.3	275	16	0805K0
10	7	14.3	286	17	09HRIA
11	7	14.3	317	16	09ISCA
12	7	14.3	387	7	096H87
13	7	14.3	467	12	0918H2
14	7	14.3	512	12	091CJ5
15	7	14.3	523	11	091W47
16	7	14.3	573	12	080S15

17	7	14.3	660	16	08YP57	08YP57 anabaena sp
18	7	14.3	4767	5	017301	017301 caenorhabdi
19	6	12.2	56	12	088224	088224 sin nombre
20	6	12.2	56	12	082160	082160 unidentified
21	6	12.2	95	11	099K39	099K39 mus musculu
22	6	12.2	99	10	08R2M7	08R2M7 oryza sativ
23	6	12.2	100	5	027316	027316 drosophila
24	6	12.2	100	5	09VUS3	09VUS3 drosophila
25	6	12.2	103	9	09GOM7	09GOM7 roseophage
26	6	12.2	132	5	P90849	P90849 caenorhabdi
27	6	12.2	136	12	041032	041032 parametium
28	6	12.2	143	10	041658	041658 vicia faba
29	6	12.2	144	16	0981C6	0981C6 rhizobium 1
30	6	12.2	166	2	087779	087779 mycobacteri
31	6	12.2	169	2	047896	047896 tolypotrixi
32	6	12.2	185	16	082U02	082U02 salmonella
33	6	12.2	185	16	0820S9	0820S9 salmonella
34	6	12.2	198	2	09F419	09F419 mycobacteri
35	6	12.2	210	16	09K7X3	09K7X3 bacillus ha
36	6	12.2	214	16	08YV80	08YV80 anabaena sp
37	6	12.2	217	16	08Y1H6	08Y1H6 ralstonia s
38	6	12.2	221	16	08Y1B2	08Y1B2 ralstonia s
39	6	12.2	231	2	007473	007473 staphylococ
40	6	12.2	234	2	093HB2	093HB2 streptomyce
41	6	12.2	235	16	09K0X9	09K0X9 neisseria m
42	6	12.2	242	16	09JTM0	09JTM0 neisseria m
43	6	12.2	248	5	09VXK4	09VXK4 drosophila
44	6	12.2	252	16	08YJF8	08YJF8 bruceella me
45	6	12.2	264	16	088025	088025 streptomyce

## ALIGNMENTS

RESULT 1  
ID 09TT28 PRELIMINARY; PRT; 34 AA.

AC 09TT28;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DT Bone gamma-carboxyglutamate protein osteocalcin (Fragment).  
GN GLA.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPRAIN-BREED BEAGLE;  
RA Wezeman F.H., Moskal S.F. II.;  
RT "Canine Osteoblast Gene Expression.";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF205942; AAF18456.1; -;  
DR InterPro: IPR002384; GLA.bone.  
DR InterPro: IPR000294; VitK\_dep\_GLA.  
DR Pfam: PF00594; GLA.1.  
DR PRINTS; PR00002; GLABONE.  
DR SMART; SM00069; GLA.1.  
FT NON\_TER 1  
FT NON\_TER 34  
SQ SEQUENCE 34 AA; 3807 MW; 3F323F8D6FDC65D8 CRC64;

Query Match 26.5%; Score 13; DB 6; Length 34;  
Best local Similarity 100.0%; Pred. No. 4.8e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELADHIGFOEA 41  
Db 22 CDELADHIGFOEA 34

RESULT 2

0925S3  
ID 0925S3 PRELIMINARY; PRT; 147 AA.  
AC 0925S3;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE MRP3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RX PubMed=11819679;  
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X., Su C.;  
RT "Mechanism of exogenous nucleic acids and their precursors improving the repair of intestinal epithelium after irradiation in mice.";  
RL World J. Gastroenterol. 6:709-717(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RA Cui D., Zeng G., Yan X., Li X., Su C.;  
RT "Cloning of mouse genes related to repairing of intestinal epithelium of the irradiated mice by treatment with the intestinal RNA of mice of the same strain.";  
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).  
DR EMBL; AF240166; AK43731.1; -  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig\_1.  
SQ SEQUENCE 147 AA; 16274 MW; 800594A12B97191F CRC64;

Query Match 26.5%; Score 13; DB 11; Length 147;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAVPYPPDLEPR 19  
Db 133 GAVPYPPDLEPR 145

RESULT 3  
ID 0925S2 PRELIMINARY; PRT; 170 AA.  
AC 0925S2;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE MRP4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RX PubMed=11819679;  
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X., Su C.;  
RT "Mechanism of exogenous nucleic acids and their precursors improving the repair of intestinal epithelium after irradiation in mice.";  
RL World J. Gastroenterol. 6:709-717(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C;  
RA Cui D., Zeng G., Yan X., Li X., Su C.;  
RT "Cloning of mouse genes related to repairing of intestinal epithelium of the irradiated mice by treatment with the intestinal RNA of mice of the same strain.";  
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).  
DR EMBL; AF240167; AK43732.1; -  
DR InterPro; IPR003598; Ig\_C2.

DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig\_1.  
DR SMART; SM00408; IgC2; 1.  
KW Immunoglobulin domain.  
SQ SEQUENCE 170 AA; 17978 MW; 5042823CC6C10F38 CRC64;

Query Match 26.5%; Score 13; DB 11; Length 170;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAVPYPPDLEPR 19  
Db 156 GAVPYPPDLEPR 168

RESULT 4  
ID 090YF0 PRELIMINARY; PRT; 298 AA.  
AC 090YF0;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE CN 8 scfv.  
GN CN 8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C; TISSUE=SPLEEN;  
RX MEDLINE=20183931; PubMed=10706631;  
RA Shiohara N., Demura T., Fukuda H.;  
RT "Isolation of a vascular cell wall-specific monoclonal antibody recognizing a cell polarity by using a phage display subtraction method.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).  
DR EMBL; AB036341; BAA86833.1; -  
DR HSSP; P01607; IRET.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig\_2.  
DR SMART; SM00406; IgV; 2.  
SQ SEQUENCE 298 AA; 31867 MW; E0F9688A17004317 CRC64;

Query Match 26.5%; Score 13; DB 11; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAVPYPPDLEPR 19  
Db 284 GAVPYPPDLEPR 296

RESULT 5  
ID 09XE33 PRELIMINARY; PRT; 205 AA.  
AC 09XE33;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Similar to sequence of BAC F7G19 from Arabidopsis thaliana.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(Gas) genomic DNA, chromosome 8, PAC clone:P0026F07.";

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP000364; BAA81759.1; -  
DR InterPro: IPR004822; Histone\_core.  
DR InterPro: IPR002965; P\_rich\_extensions.  
DR PRINTS: PR01217; PRICHEXTENSN.  
SQ SEQUENCE 205 AA; 21909 MW; F747D35F86B59C8 CRC64;

Query Match 16.3%; Score 8; DB 10; Length 205;  
Best Local Similarity 100.0%; Pred. No. 0.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVPP 13  
|||||||  
Db 17 LGAPVPP 24

RESULT 6  
039303 PRELIMINARY; PRT; 536 AA.

AC 039303;  
DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JUN-1998 (TREMblrel. 05, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Counterpart of HSV-1 gene RL2 and VZV gene 61.  
GN 63.  
OS Equine herpesvirus 4.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicelloviruses.  
OX NCBI\_TaxID=10331;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NS80567;  
RX MEDLINE=98264497; PubMed=9603335.  
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;  
RT "The DNA sequence of equine herpesvirus-4.";  
RL J. Gen. Virol. 79:1197-1203(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NS80567;  
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL: AF030027; AAC59582.1; -  
DR HSSP: P28990; 1CHC.  
DR InterPro: IPR001841; ZnF\_fing.  
DR Pfam: PF00097; zf-C3HC4; 1.  
DR SMART: SM00184; RING; 1.  
DR PROSITE: PS00518; zf\_RING\_1; 1.  
KW Zinc-finger.  
SQ SEQUENCE 536 AA; 59686 MW; A973B9B2A92DD08 CRC64;

Query Match 16.3%; Score 8; DB 12; Length 536;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PVYPDPL 16  
|||||||  
Db 486 PVYPDPL 493

RESULT 7  
090VM2 PRELIMINARY; PRT; 97 AA.

AC 090VM2;  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE Osteocalcin.  
OS Sparus aurata (Gilthead sea bream).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
OC Sparidae; Sparus.

OX NCBI\_TaxID=8175;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21297182; PubMed=11404005;  
RA Plato J.P., Ohresser M.C.P., Cancela M.L.;  
RT "Cloning of the bone gla protein gene from the teleost fish Sparus  
aurata. Evidence for overall conservation in gene organization and  
RT bone-specific expression from fish to man.";  
RL Gene 270:77-91(2001).  
DR EMBL: AF048703; AAK6568.1; -  
DR InterPro: IPR000294; Vitk\_dep\_GLA.  
DR Pfam: PF00594; gla; 1  
DR PROSITE: PS00011; GLU\_CARBOXYLATION; UNKNOWN\_1.  
SQ SEQUENCE 97 AA; 10434 MW; B95608824FDEECB CRC64;

Query Match 14.3%; Score 7; DB 13; Length 97;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 REVCENL 26  
|||||||  
Db 66 REVCENL 72

RESULT 8  
069068 PRELIMINARY; PRT; 268 AA.

AC 069068;  
DT 01-AUG-1998 (TREMblrel. 07, Created)  
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE HX1 protein (Fragment).  
GN HX1.  
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NM88;  
RX MEDLINE=99008986; PubMed=9791102;  
RA Metcalf W.W., Wolfe R.S.;  
RT "Molecular genetic analysis of phosphite and hypophosphite oxidation  
by Pseudomonas stutzeri NM88.";  
RL J. Bacteriol. 180:5547-5558(1998).  
CC -1- FUNCTION: BELONGS TO AN OPERON INVOLVED IN HYPOPHOSPHITE  
OXIDATION.  
CC -1- SIMILARITY: BELONGS TO THE PHN1 FAMILY.  
DR EMBL: AF061267; AAC71719.1; -  
FT NON\_TER 268 268  
SQ SEQUENCE 268 AA; 30070 MW; C4ADA5E14C9BDB63 CRC64;

Query Match 14.3%; Score 7; DB 2; Length 268;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VYPDPL 16  
|||||||  
Db 133 VYPDPL 139

RESULT 9  
0805K0 PRELIMINARY; PRT; 275 AA.

AC 0805K0;  
DT 01-JUN-2002 (TREMblrel. 21, Created)  
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE AGR\_C\_880P.  
GN AGR\_C\_880.  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

```

OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gatlung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlet K., Gordon J., Vaubin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL: AE007985; AAK6310.1;
SQ SEQUENCE 275 AA; 29911 MW; CFI40634E33C0322 CRC64;

Query Match 14.3%; Score 7; DB 16; Length 275;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 FQAYRR 44
DB 167 FQAYRR 173

RESULT 10
O9HRI4 PRELIMINARY; PRT; 286 AA.
AC O9HRI4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE 3-hydroxyacyl-CoA dehydrogenase.
GN HBDI OR VMG0681G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leitner B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angewine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL: AE005014; AAG19174.1;
DR HSP; P00348; 3HSD.
DR InterPro: IPR002135; 3HSDH.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF00725; 3HSDH_1.
DR Pfam: PF02737; 3HSDH_N; 1.
KW Complete proteome.
SQ SEQUENCE 286 AA; 30687 MW; 880F72A7D3AB2342 CRC64;

Query Match 14.3%; Score 7; DB 17; Length 286;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHIG 37
DB 234 ELADHIG 240

RESULT 11
O915C4 PRELIMINARY; PRT; 317 AA.
ID O915C4

```

```

AC O915C4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Probable transcriptional regulator.
GN PA0815.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stoyer C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- SIMILARITY: TO THE LYSR FAMILY OF TRANSCRIPTIONAL REGULATORS.
DR EMBL: AE004516; AAG04204.1;
DR InterPro: IPR000847; HTH_LysR.
DR InterPro: IPR005119; LysR_subst.
DR Pfam: PF00126; HTH_1; 1
DR PRINTS: PR00039; HTH_LysR.
DR Pfam: PF03466; LysR_substrate; 1.
DR PROSITE: PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
KW DNA-binding; Transcription regulation; Complete proteome.
SQ SEQUENCE 317 AA; 35990 MW; CC73FDB4A3473960 CRC64;

Query Match 14.3%; Score 7; DB 16; Length 317;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 OMLGAPV 10
DB 42 OMLGAPV 48

RESULT 12
O96H87 PRELIMINARY; PRT; 387 AA.
AC O96H87;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:10474).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: BC008819; AAH08819.1;
DR InterPro: IPR000923; BlueCu_1.
DR InterPro: IPR000536; Hormone_rec_1lg.
DR InterPro: IPR000822; ZnF_C2H2.
DR InterPro: IPR001628; ZnF_C4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR ProDom: PD000035; ZnF_C4steroid; 1.
DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; UNKNOWN_1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.

```



KW DNA-binding: Nuclear protein; Receptor; Transcription regulation;  
 KW Zinc-finger.

SO SEQUENCE 387 AA; 43555 MW; 9C96CF2B8E66403C CRC64;

Query Match 14.3%; Score 7; DB 4; Length 387;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WLGAPVP 11

DB 4 WLGAPVP 10

RESULT 13

O918H2 ID O918H2 PRELIMINARY; PRT; 467 AA.

AC O918H2; 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE B84.

OS baboon cytomegalovirus.

OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC betaherpesvirinae; Cytomegalovirus.

OX NCBI\_TaxID=120505;

RN [1] SEQUENCE FROM N.A.

RC STRAIN-OCOM4-37;

RA Blewett E., Rogers P., Kravitz R., Barry P.;

RT "The UL82 gene families of baboon (BacMV) and rhesus cytomegalovirus

RT (rhesus cytomegalovirus) each contain 4 genes and are closely related

RT to the UL82 family (3 genes) of human cytomegalovirus.";

RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL:AA411694; AAL10297.1; -

DR InterPro: IPR000637; AT\_hook.

DR Pfam: PF02178; AT\_hook.1.

SO SEQUENCE 467 AA; 52651 MW; 335A9A9FA4831CC CRC64;

Query Match 14.3%; Score 7; DB 12; Length 467;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYGP 48

DB 333 YRRFYGP 339

RESULT 14

O91CJ5 ID O91CJ5 PRELIMINARY; PRT; 512 AA.

AC O91CJ5; 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE R84.

OS Rhesus cytomegalovirus (strain 68-1) (RhCMV).

OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC betaherpesvirinae; Cytomegalovirus.

OX NCBI\_TaxID=103930;

RN [1] SEQUENCE FROM N.A.

RC STRAIN-68-1;

RA Blewett E.L., Preston R.P., Kravitz R., Barry P.;

RT "The UL82 gene families of baboon (baboon cytomegalovirus) and rhesus

RT cytomegalovirus (RhCMV) each contain 4 genes and are closely related

RT to the UL82 family (3 genes) of HCMV.";

RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.

DR [2] SEQUENCE FROM N.A.

DR STRAIN-68-1;

RA Kravitz R.H., Barry P.A.;

RT "Simian cytomegaloviruses as models for HCMV persistence and  
 RT pathogenesis.";

RT Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL:AF078698; AAL08223.1; -

SO SEQUENCE 512 AA; 57341 MW; EF18BA8354A39656 CRC64;

Query Match 14.3%; Score 7; DB 12; Length 512;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYGP 48

DB 367 YRRFYGP 373

RESULT 15

O91W47 ID O91W47 PRELIMINARY; PRT; 523 AA.

AC O91W47; 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical 57.3 kDa protein.

GN AW743111.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1] SEQUENCE FROM N.A.

RC TISSUE-EYE, AND RETINA;

RA Strassberg R.;

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL:BC017129; AAH17129.1; -

DR MGD; MG1:2144585; AW743111.

KW Hypothetical protein.

SO SEQUENCE 523 AA; 57296 MW; 7AEAB121D7BF4009 CRC64;

Query Match 14.3%; Score 7; DB 11; Length 523;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LEPRREV 22

DB 217 LEPRREV 223

RESULT 16

O80S15 ID O80S15 PRELIMINARY; PRT; 573 AA.

AC O80S15; 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE UL84.

OS chimpanzee cytomegalovirus.

OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC betaherpesvirinae; Cytomegalovirus.

OX NCBI\_TaxID=188763;

RN [1] SEQUENCE FROM N.A.

RA Davison A.J., Akter P., Dolan A., Wright K.M., Addison C.;

RA Alencor D.J., Hayward G.S., McGeoch D.J.;

RT "The human cytomegalovirus genome revisited.";

RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL:AF480884; AAM00723.1; -

SO SEQUENCE 573 AA; 63004 MW; B56C2AD35F76D020 CRC64;

Query Match 14.3%; Score 7; DB 12; Length 573;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYGP 48

```

Db      406 YRRFYGP 412

|||||
RESULT 17
08P57      PRELIMINARY;      PRT;      660 AA.
AC      08P57;
DT      01-MAR-2002 (TREMBLrel. 20, Created)
DT      01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Hypothetical protein A114343.
GN      AL14343.
OS      Anabaena sp. (strain PCC 7120).
OC      Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX      NCBI_TaxID=103690;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=21595285; PubMed=11759840;
RA      Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA      Mamanabe A., Iriuchimi M., Ishikawa A., Kawashima K., Kimura T.,
RA      Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA      Nakazaki N., Shimp S., Sugimoto M., Takazawa M., Yamada M.,
RA      Yasuda M., Tabata S.;
RT      "Complete genomic sequence of the filamentous nitrogen-fixing
RT      cyanobacterium Anabaena sp. strain PCC 7120."
RL      DNA Res. 8:205-213(2001).
DR      EMBL: AP003596; BAB76042.1; -.
DR      InterPro: IPR000901; CPsase.
DR      PROSITE: PS00867; CPsase_2; UNKNOWN_1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 660 AA; 71793 MW; B85832335E20472 CRC64;

Query Match      14.3%; Score 7; DB 16; Length 660;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      13 PDPLEPR 19
|||||
Db      312 PDPLEPR 318

RESULT 18
017301      PRELIMINARY;      PRT;      4767 AA.
AC      017301;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      G01D9.5 protein.
GN      G01D9.5.
OS      Caenorhabditis briggsae.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC      Rhabditidae; Pelodierinae; Caenorhabditis.
OX      NCBI_TaxID=6238;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN-GUTARAT G16;
RA      Wu X, Le TT.;
RT      "The sequence of C. briggsae cosmid G01D9."
RT      Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      STRAIN-GUTARAT G16;
RA      Waterston R.;
RT      "The C. briggsae Genome Sequencing Project."
RT      Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL: U56248; AAA98699.1; -.
DR      HSSP: P14687; IAMU.
DR      InterPro: IPR001227; Ac.transferase.
DR      InterPro: IPR000873; AMP-bind.
DR      InterPro: IPR001242; Condensatn.
DR      InterPro: IPR000794; Ketoacyl-synt.

```

```

DR      InterPro: IPR000734; Lipase.
DR      InterPro: IPR003880; Pantine_attach.
DR      InterPro: IPR000379; Ser.estrs.site.
DR      InterPro: IPR000130; Zn.MRpeptidse.
DR      Pfam: PF00698; Acyl.transf. 2.
DR      Pfam: PF00501; AMP-binding; 1.
DR      Pfam: PF00668; Condensation; 1.
DR      Pfam: PF00109; Ketoacyl-synt; 2.
DR      Pfam: PF02801; ketoacyl-synt_C; 2.
DR      Pfam: PF00550; pp-binding; 5.
DR      PROSITE: PS50075; ACP_DOMAIN; 5.
DR      PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
DR      PROSITE: PS00120; LIPASE_SER; UNKNOWN_1.
DR      PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_2.
DR      PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW      Phosphopantetheine.
SQ      SEQUENCE 4767 AA; 535178 MW; 0499BB847CB7A07A CRC64;

Query Match      14.3%; Score 7; DB 5; Length 4767;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      30 DELADHI 36
|||||
Db      35 DELADHI 41

RESULT 19
088224      PRELIMINARY;      PRT;      56 AA.
AC      088224;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      Nucleocapsid protein (Fragment).
OS      Sin Nombre virus.
OC      Viruses; ssRNA negative-strand viruses; Bunyaviridae; Hantavirus.
OX      NCBI_TaxID=37705;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN-NV WA-R199;
RA      MEDLINE=96036486; PubMed=7483255;
RA      Rowe J.E., St Jeor S.C., Riolo J., Oateson E.W., Monroe M.C.,
RA      Henderson W.W., Ksiazek T.G., Rollin P.E., Nichol S.T.;
RT      "Coexistence of several novel hantaviruses in rodents indigenous to
RT      North America."
RT      Virology 213:122-130(1995).
DR      EMBL: U33259; AAC36797.1; -.
DR      InterPro: IPR002214; Hanta_nucleocap.
DR      Pfam: PF00846; Hanta_nucleocap; 1.
DR      ProDom: PD001501; Hanta_nucleocap; 1.
FT      NON_TER      1
FT      NON_TER      1
SQ      SEQUENCE 56 AA; 6107 MW; 15BE701F402F7D3B CRC64;

Query Match      12.2%; Score 6; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      31 ELADHI 36
|||||
Db      23 ELADHI 28

RESULT 20
082160      PRELIMINARY;      PRT;      56 AA.
AC      082160;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      Nucleocapsid protein (Fragment).
OS      unidentified.

```

```

OC unclassified.
OX NCBI_TaxID=32644;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=STN NOMBRE/CA MO-R159, AND STN NOMBRE/NV NY-R301;
RX MEDLINE=96130200; PubMed=8553562;
RA Henderson W.W., Monroe M.C., St Jeor S.C., Thayer W.P., Rowe J.E.,
  Peters C.J., Nichol S.T.;
RT "Naturally occurring Stn NOMBRE virus genetic reassortants.";
RL Virology 214:602-610(1995).
DR EMBL: U45015; AAB48145.1; -.
DR EMBL: U45016; AAB48146.1; -.
DR InterPro: IPR002214; Hanta_nucleocap.
DR Pfam: PF00846; Hanta_nucleocap; 1.
DR Prodom: PD001501; Hanta_nucleocap; 1.
KW Nucleocapsid.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6107 MW; 15BE701F402F7D3B CRC64;

Query Match
Best Local Similarity 12.2%; Score 6; DB 12; Length 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHI 36
DB 23 ELADHI 28

RESULT 21
Q99K39 PRELIMINARY; PRT; 95 AA.
AC Q99K39;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Similar to bone gamma-carboxylglutamate protein, related sequence
  1.
GN A1461847.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC005483; AAB05483.1; -.
DR MGD: MGI:2139729; A1461847.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; Vitr_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU CARBOXYLATION; 1.
SQ SEQUENCE 95 AA; 10445 MW; 25C4B3A6B51909E CRC64;

Query Match
Best Local Similarity 12.2%; Score 6; DB 11; Length 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 PDPLEP 18
DB 58 PDPLEP 63

RESULT 22
O8RZM7 PRELIMINARY; PRT; 99 AA.
AC O8RZM7;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

```

```

DE B1168H06.7 protein.
GN B1168H06.7.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC
  clone:B1168H06.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003563; BAB89169.1; -.
SQ SEQUENCE 99 AA; 10491 MW; 257199E32438B5E6 CRC64;

Query Match
Best Local Similarity 12.2%; Score 6; DB 10; Length 99;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11
DB 25 LGAPVP 30

RESULT 23
Q27316 PRELIMINARY; PRT; 100 AA.
AC Q27316;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE L71-1.
GN EIG71EA OR L71-1 OR CG16931.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriodea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R; TISSUE=SALIVARY GLAND;
RA Wright L.G., Chen T., Thummel C.S., Guild G.M.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Wright L.G., Chen T., Thummel C.S., Guild G.M.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U24095; AAA65109.1; -.
DR EMBL: U23836; AAA74176.1; -.
DR FLYBase: FBgn0004588; EIG71EA.
DR InterPro: IPR003475; Insect_Unk.
DR Pfam: PF02448; L71; 1.
SQ SEQUENCE 100 AA; 11895 MW; 0D7A22A0639D38A8 CRC64;

Query Match
Best Local Similarity 12.2%; Score 6; DB 5; Length 100;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 DCDELA 33
DB 24 DCDELA 29

RESULT 24
Q9VUS3 PRELIMINARY; PRT; 100 AA.
AC Q9VUS3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

```

```

DE EIG71EA protein.
GN EIG71EA OR CG16931.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN NCB1_Taxid=7227;
RX
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratilake P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.L.G.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandal D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Flatschmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.Y., Moadary C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003530; AAF49602.1;
DR FlyBase: FBgn004588; EIG71EA.
DR InterPro: IPR003475; Insect_unk.
DR Pfam: PF02448; L71.1.
SQ SEQUENCE 100 AA; 11867 MW; 1BD34D1048B5FA8 CRC64;

```

Query Match 12.2%; Score 6; DB 5; Length 100;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 28 DCDELA 33
DB 24 DCDELA 29

```

RESULT 25  
O9G0H7  
AC O9G0H7; PRELIMINARY; PRT; 103 AA.  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Gp8.  
OS Roseophage SI01.

```

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCB1_Taxid=136084;
RN [1]
RP SEQUENCE FROM N.A.
RA Rohwer F.L., Segall A.M., Steward G., Seguritan V., Breitbart M.,
RA Wolven F., Azam F.;
RT "The complete genomic sequence of the marine phage Roseophage SI01
RT shares homology with nonmarine phages."
RL J. Virol. 74:408-418(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Rohwer F.L., Segall A.M., Steward G., Seguritan V., Breitbart M.,
RA Wolven F., Azam F.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF189021; AAC02591.1;
SQ SEQUENCE 103 AA; 11521 MW; DEF26AAB23CD725 CRC64;

```

Query Match 12.2%; Score 6; DB 9; Length 103;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 24 ELPNDC 29
DB 12 ELPNDC 17

```

RESULT 26  
P90849  
ID P90849; PRELIMINARY; PRT; 132 AA.  
AC P90849;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE F26E4.9 protein.  
GN F26E4.9  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCB1\_Taxid=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA lightning J.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C. elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL: Z81070; CAB03002.1;
DR InterPro: IPR002124; COX5B.
DR Pfam: PF01215; COX5B.1.
DR ProDom: PD007270; COX5B.1.
SQ SEQUENCE 132 AA; 14733 MW; 6D997BEB8626554A CRC64;

Query Match 12.2%; Score 6; DB 5; Length 132;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 12 YPDPLE 17
DB 39 YPDPLE 44

```

RESULT 27  
O41032  
ID O41032; PRELIMINARY; PRT; 136 AA.  
AC O41032;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE A550R protein.  
 GN A550R.  
 OS Paramesicium bursaria chloroella virus 1 (PBCV-1).  
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.  
 NC NCB1\_TaxID=10506;  
 RX MEDLINE-98022962; PubMed-9356347;  
 RA Li Y., Lu Z., Sun L., Ropp S., Kutish G.F., Rock D.L., Van Etten J.L.;  
 RT "Analysis of 74 kb of DNA located at the right end of the 330-kb  
 RT chloroella virus PBCV-1 genome.";  
 RL Virology 237:360-377(1997).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20013326; PubMed-10544099;  
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,  
 RA Lisec A.D., Nickerson K.W., Van Etten J.L.;  
 RT "Chloroella virus PBCV-1 encodes a functional homospermidine  
 RT synthase.";  
 RL Virology 263:254-262(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20478054; PubMed-11021991;  
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;  
 RT "Characterization of a beta-1,3-glucanase encoded by chloroella virus  
 RT PBCV-1.";  
 RL Virology 276:27-36(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Van Etten J.L.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Van Etten J.L.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Van Etten J.L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Graves M.V., Van Etten J.L.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RA Graves M.V., Van Etten J.L.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RA Gurnon J.R., Graves M.V., Van Etten J.L.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U42580; AAC96997.1; -;  
 SQ SEQUENCE 136 AA; 1556 MW; 18AFC2D446840481 CRC64;  
 Query Match 12.2%; Score 6; DB 12; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 LGAPVP 11  
 Db 12 LGAPVP 17

RESULT 28  
 ID 041658 PRELIMINARY; PRT; 143 AA.  
 AC 041658;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE MADH ubiquinone oxidoreductase subunit 5 (nads).  
 OS Vicia faba (Broad bean).  
 OG Mitochondrion.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosid I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.  
 NC NCB1\_TaxID=3906;  
 RX MEDLINE-10544099; PubMed-10544099;  
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,  
 RA Lisec A.D., Nickerson K.W., Van Etten J.L.;  
 RT "Chloroella virus PBCV-1 encodes a functional homospermidine  
 RT synthase.";  
 RL Virology 263:254-262(1999).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20013326; PubMed-10544099;  
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,  
 RA Lisec A.D., Nickerson K.W., Van Etten J.L.;  
 RT "Chloroella virus PBCV-1 encodes a functional homospermidine  
 RT synthase.";  
 RL Virology 263:254-262(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20478054; PubMed-11021991;  
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;  
 RT "Characterization of a beta-1,3-glucanase encoded by chloroella virus  
 RT PBCV-1.";  
 RL Virology 276:27-36(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Van Etten J.L.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Van Etten J.L.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Van Etten J.L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Graves M.V., Van Etten J.L.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RA Graves M.V., Van Etten J.L.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RA Gurnon J.R., Graves M.V., Van Etten J.L.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U42580; AAC96997.1; -;  
 SQ SEQUENCE 144 AA; 16083 MW; EBDFADEC8ADBADP3 CRC64;  
 Query Match 12.2%; Score 6; DB 10; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 APVPVP 13  
 Db 55 APVPVP 60

RESULT 30  
 ID 087779 PRELIMINARY; PRT; 166 AA.  
 AC 087779;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical 18.1 kDa protein (Fragment).  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.

```

OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19698;
RX MEDLINE=93328703; PubMed=8335649;
RA Gilot P., De Kessel M., Coene M., Machetlinckx L., Cocito C.;
RT "Isolation and sequencing of the gene coding for an antigenic 34-
RL kilodalton protein of Mycobacterium paratuberculosis.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19698;
RX MEDLINE=93159670; PubMed=10068253;
RA Gilot P.;
RT "Specificity of the 34-kilodalton immunodominant protein of
RL Mycobacterium avium subsp. paratuberculosis.";
DR Clin. Diagn. Lab. Immunol. 6:146-148(1999).
EMBL: X68102; CAA48222.1; -.
DR InterPro: IPR002103; Bac_Luciferase.
DR Pfam: PF00296; Bac_Luciferase; 1.
KW Hypothetical protein.
FT NON_TER 166
SQ SEQUENCE 166 AA; 18052 MW; B7A8F0F7A6A5344 CRC64;

Query Match 12.2%; Score 6; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 DELADH 35
DB 123 DELADH 128

RESULT 31
ID Q47896 PRELIMINARY; PRT; 169 AA.
AC Q47896;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ORF 169.
OS Tolypothrix sp. PCC 7601.
OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Tolypothrix.
OX NCBI_TaxID=1188;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88260883; PubMed=2838727;
RA Parsot C., Mazel D.;
RT "Cloning and nucleotide sequence of the thrB gene from the
RL cyanobacterium Calothrix PCC 7601.";
DR EMBL: Y00522; CAA68577.1; -.
SQ SEQUENCE 169 AA; 19182 MW; 1A036CFA6CEAA400 CRC64;

Query Match 12.2%; Score 6; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EAYRRF 45
DB 43 EAYRRF 48

RESULT 32
ID Q8ZJ02 PRELIMINARY; PRT; 185 AA.
AC Q8ZJ02;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative fibrillar subunit.
GN STD OR STM4592.
OS Salmonella typhimurium.

```

```

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Salmonella.
NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LT2 / SSGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sandersen K.E., Spleth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RL LT2.";
DR Nature 413:852-856(2001).
DR EMBL: AE008916; AL23407.1; -.
DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 185 AA; 19337 MW; C83C8701E9A9676E CRC64;

Query Match 12.2%; Score 6; DB 16; Length 185;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VPPDP 15
DB 99 VPPDP 104

RESULT 33
ID Q8Z0S9 PRELIMINARY; PRT; 185 AA.
AC Q8Z0S9;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE Putative fibrillar subunit.
GN STD OR STY4940.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Hague A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RL enterica serovar Typhi CT18.";
DR Nature 413:848-852(2001).
DR EMBL: AL627284; CAD03423.1; -.
DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 185 AA; 19425 MW; 836A2100EA432234 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 185;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 VPPDP 15
DB 99 VPPDP 104

```

RESULT 34  
 09F419 PRELIMINARY; PRT: 198 AA.  
 AC 09F419;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Putative transcription regulator.  
 GN TERR.  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20431891; PubMed=10974106;  
 RA Bull T.J., Hermon-Taylor J., Pavlik I., El-Zaatari F., Tizard M.;  
 RT "Characterization of IS900 loci in Mycobacterium avium subsp.  
 RT paratuberculosis and development of multiplex PCR typing.";  
 RL Microbiology 146:2185-2197(2000).  
 CC -1- SIMILARITY: BELONGS TO THE TERR/ACR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 DR EMBL: AJ250023; CAC10267.1; -  
 DR InterPro: IPR001647; HTH\_TERR.  
 DR Pfam: PF00440; terr; 1.  
 DR PRINTS: PR00455; HTHETR.  
 DR DNA-binding; Transcription regulation.  
 KW SEQUENCE 198 AA; 21553 MW; 06786AE1165F302 CRC64;  
 SQ

Query Match 12.2%; Score 6; DB 2; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 QEARRR 44  
 |||||  
 DB 144 QEARRR 149

RESULT 35  
 09K7X3 PRELIMINARY; PRT: 210 AA.  
 AC 09K7X3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE Acetoin dehydrogenase.  
 GN ACUA OR BH3235.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RA TAKEMI H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fujii F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL: AP001518; BAB06954.1; -  
 DR InterPro: IPR000182; GCN5acetyltransf.  
 DR Pfam: PF00583; Acetyltransf; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 210 AA; 24338 MW; 2E5242DE144EC1C7 CRC64;  
 QY

Query Match 12.2%; Score 6; DB 16; Length 210;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 12 YDPPLE 17  
 |||||

DB 80 YDPPLE 85  
 RESULT 36  
 08Y180 PRELIMINARY; PRT: 214 AA.  
 AC 08Y180;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE Hypothetical protein A114690.  
 GN A114690.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,  
 RA Matnabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120."  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL: AP003597; BAB76389.1; -  
 DR InterPro: IPR001601; Methyltransf.  
 DR Hypothetical protein; Complete proteome.  
 KW SEQUENCE 214 AA; 24688 MW; 7A4D1DC45E40F947 CRC64;  
 SQ

Query Match 12.2%; Score 6; DB 16; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 QWLGAP 9  
 |||||  
 DB 196 QWLGAP 201

RESULT 37  
 08Y186 PRELIMINARY; PRT: 217 AA.  
 AC 08Y186;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE Probable riboflavin synthase (Alpha chain) protein (EC 2.5.1.9).  
 GN RIBE OR RSC0714 OR RS05139.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM11000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoudat M., Genin S., Artiguenave F., Gouzy J., Mangnot S.,  
 RA Ariat M., Billault A., Brotlier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Sauvin W., Schlex T.,  
 RA Sigler P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646060; CAD14244.1; -  
 DR InterPro: IPR001783; Lum\_binding.  
 DR Pfam: PF00677; Lum\_binding; 2.  
 DR ProDom: PD004110; Lum\_binding; 1.  
 DR TIGRFAMS: TIGR00187; ribe; 1.  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 217 AA; 22663 MW; 44682960915B3172 CRC64;  
 QY

```

Query Match      12.2%; Score 6; DB 16; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 LADHIG 37
DB 92 LADHIG 97

RESULT 38
O8Y1B2 PRELIMINARY; PRT; 221 AA.
AC O8Y1B2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Hypothetical protein Rsc0778.
GN RSC0778 OR RS05076.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cottolico L.,
RA Chandier M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Levy W., Moisan A., Robert C., Saulin W., Schlex T.,
RA Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AL646061; CAD14480.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 221 AA; 24486 MW; A9EB517D9402E246 CRC64;

Query Match      12.2%; Score 6; DB 16; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11
DB 32 LGAPVP 37

RESULT 39
O07473 PRELIMINARY; PRT; 231 AA.
AC O07473;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE GdmF.
GN GDMF.
OS Staphylococcus gallinarum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1293;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUE3928;
RX MEDLINE=89306540; PubMed=2765032;
RA Schnell N., Entian K.D., Gotz F., Horner T., Kellner R., Jung G.;
RT "Structural gene isolation and prepeptide sequence of gallidermin, a
RT new lantibiotic containing antibiotic.";
RL FEWS Microbiol. Lett. 49:263-267(1989).
GN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=TUE3928;
RC STRAIN=TUE3928;
RX MEDLINE=97294510; PubMed=9150266;
RA Peschel A., Schnell N., Hille M., Entian K.D., Gotz F.;
RN [1]

```

```

RT "Secretion of the lantibiotics epidermin and gallidermin: sequence
RT analysis of the genes gdmT and gdmH, their influence on epidermin
RT production and their regulation by EpiQ.";
RL Mol. Gen. Genet. 254:312-318(1997).
DR EMBL: D61158; AAB61132.1; -
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_Transport.
DR Pfam: PF00005; ABC_tran; 1.
DR ProDom: PD000006; ABC_Transport; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW ATP-binding.
SQ SEQUENCE 231 AA; 25902 MW; 2475D9B3F5D237A0 CRC64;

Query Match      12.2%; Score 6; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 LADHIG 37
DB 195 LADHIG 200

RESULT 40
O93HB2 PRELIMINARY; PRT; 234 AA.
AC O93HB2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE CIPX homolog.
GN CIPX.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kiuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: Deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL: AB070947; BAB69277.1; -
DR InterPro: IPR004176; CIP_N.
DR Pfam: PF02861; CIP_N; 2.
SQ SEQUENCE 234 AA; 24300 MW; AFC1EF7EB85C4B7C CRC64;

Query Match      12.2%; Score 6; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 LADHIG 37
DB 34 LADHIG 39

RESULT 41
O9K0X9 PRELIMINARY; PRT; 235 AA.
AC O9K0X9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cell division protein FtsQ.
GN NMB0425.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]

```



RP SEQUENCE FROM N.A.  
 RC STRAIN-MC58 / SEROGROUP B;  
 RX MEDLINE-20175755; PubMed-10710307;  
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
 RA Eissen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
 RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,  
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Cifton H., Clark E.B.,  
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Yamaharan J.,  
 RA Gail J., Scarlato V., Maignani V., Pizsa M., Grandi G., Sun L.,  
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;  
 RT \*Complete genome sequence of *Neisseria meningitidis* serogroup B strain  
 MC58.";  
 RL Science 287:1809-1815(2000).  
 DR EMBL: AE002398; AAF40863.1; -;  
 DR TIGR: NMB0425; -;  
 KW Complete proteome.  
 SQ SEQUENCE 235 AA; 27135 MW; B7E605926BD5E9FF CRC64;

Query Match 12.2%; Score 6; DB 16; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 39 QEAYR 44  
 Db 70 QEAYR 75

RESULT 42  
 ID 09JTO0 PRELIMINARY; PRT; 242 AA.  
 AC 09JTO0;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Cell division protein.  
 GN FTSQ OR NMA2059.  
 OS *Neisseria meningitidis* (serogroup A).  
 OC Bacteria; Proteobacteria; beta subdivision; *Neisseriaceae*; *Neisseria*.  
 OX NCBI\_TaxID=56599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-22491 / SEROGROUP A / SEROTYPE 4A;  
 RX MEDLINE-2022556; PubMed-10761919;  
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,  
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,  
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,  
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
 RA Whitehead S., Spratt B.G., Barrrell B.G.;  
 RT \*Complete DNA sequence of a serogroup A strain of *Neisseria*  
 RT meningitidis 22491.";  
 RL Nature 404:502-506(2000).  
 DR EMBL: AL162758; CAB85277.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 242 AA; 27933 MW; A29018736D2FE35 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 39 QEAYR 44  
 Db 77 QEAYR 82

RESULT 43  
 ID 09YX4 PRELIMINARY; PRT; 248 AA.  
 AC 09YX4;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)

DE CG9066 protein (LD12946P).  
 GN CG9066.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY.  
 RX MEDLINE-2019606; PubMed-10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Adair J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyama C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinelt K., Remington K., Saunders R.D.C., Scheer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT \*The genome sequence of *Drosophila melanogaster*.  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY.  
 RA Stapleton M., Broksstein P., Hong L., Abbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunco J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003500; AAF48534.1; -;  
 DR EMBL: AY061163; AL28711.1; -;  
 DR FLYBase: FBgn0030703; CG9066.  
 DR InterPro: IPR001199; Cyt\_B5.  
 DR pfam: PF001173; heme\_1.1.  
 SQ SEQUENCE 248 AA; 27921 MW; 1A7B9C7BDE72FA3 CRC64;

Query Match 12.2%; Score 6; DB 5; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 43 RREYGP 48  
 Db 115 RREYGP 120

## RESULT 44

08YJF8 PRELIMINARY: PRT: 252 AA.  
 AC 08YJF8:  
 DT 01-MAR-2002 (TREMblrel. 20, Created)  
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Acetyltransferase (EC 2.3.1.-).  
 GN BME0125.  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA Delvecchio V.G., Kapatali V., Redkar R.J., Patra G., Mujer C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,  
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Haselkorn R., Kyrpides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
 DR EMBL: AE009455; AL51307.1: -;  
 DR InterPro: IPR00182; GCN5acetyltransf.  
 DR Pfam: PF00583; Acetyltransferase; 1.  
 KW Transferase; Acyltransferase; Complete proteome.  
 SQ SEQUENCE 252 AA; 27881 MW; 5A9582613939E773 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 252;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 37 GEOEAY 42  
 |||||  
 DB 237 GEOEAY 242

## RESULT 45

088025 PRELIMINARY: PRT: 264 AA.  
 AC 088025:  
 DT 01-NOV-1998 (TREMblrel. 08, Created)  
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Putative transcriptional regulator.  
 GN SCO6669 OR SC5A7.19C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).

[4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -1 SIMILARITY: BELONGS TO THE ICLR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC EMBL: AL031107; CA19948.1: -;  
 DR InterPro: IPR00285; HTH\_ICLR.  
 DR Pfam: PF01614; ICLR; 1.  
 DR SMART: SM00346; HTH\_ICLR; 1.  
 KW DNA-binding; Transcription regulation.  
 SQ SEQUENCE 264 AA; 28066 MW; 99C0F97B015D3709 CRC64;

Query Match 12.2%; Score 6; DB 16; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 FOEAYR 43  
 |||||  
 DB 153 FOEAYR 158

Search completed: December 4, 2002, 15:39:37  
 Job time : 32 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:40:10 ; Search time 36 Seconds  
(without alignments)  
181.369 Million cell updates/sec

```
Title: US-09-462-931-2-COPY
Perfect score: 272
Sequence: 1 YLYQWLGAPVPYPDDLXPRR.....DELADHIGFQEA YRRFYGPV 49
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : A\_Geneseq\_010002.\*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	266	97.8	49	14	AAK32936	21, 24, Gla human
2	266	97.8	49	14	AAK32937	17, 21, 24, Gla h
3	266	97.8	49	18	AAW34263	Glul7-osteocalcin
4	266	97.8	49	18	AAW34264	Glul7-osteocalcin
5	266	97.8	49	18	AAW34265	Gla21-osteocalcin
6	266	97.8	49	18	AAW34266	Gla21-osteocalcin
7	266	97.8	49	20	AAW9811	Human osteocalcin
8	266	97.8	51	12	AAK10146	Human osteocalcin
9	266	97.8	98	19	AAW6094	Human osteocalcin
10	266	97.8	100	12	AAK10147	Human osteocalcin

1.1	266	97.8	100	2.3	AAK00687	Reference sequence
1.2	259	95.2	97	1.3	AAK20043	Fusion protein for
1.3	258	94.9	49	1.6	AAK01681	Bone Gla protein.
1.4	258	94.9	100	2.5	AAK01688	Bone Gla protein.
1.5	247	90.8	47	2.2	AAK01917	Polymorphic variant
1.6	237.5	87.3	46	2.2	AAK01919	Bone Gla protein p
1.7	202	74.3	99	1.9	AAK076097	Rat osteocalcin p
1.8	193	71.0	49	2.3	AAK05975	Modified osteocalc
1.9	162	59.6	38	8	AAK17282	Pepide from mamma
2.0	148	54.4	95	1.9	AAK76095	Mouse osteocalcin
2.1	144	52.9	95	1.9	AAK76096	Mouse osteocalcin
2.2	113	41.5	21	1.1	AAK06630	Human osteocalcin
2.3	108	39.7	20	1.3	AAK26407	Osteocalcin AA 1-1
2.4	81.5	30.0	68	2.3	AAK17841	Pepide presentatn
2.5	77	28.3	15	1.1	AAK06631	Human osteocalcin
2.6	74	27.2	38	2.3	AAK08451	Insulin/insulin-11
2.7	74	27.2	39	2.3	AAK08442	Insulin/insulin-11
2.8	74	27.2	55	2.3	AAK08452	Insulin/insulin-11
2.9	74	27.2	55	2.3	AAK08497	Insulin/insulin-11
3.0	74	27.2	58	2.3	AAK08459	Insulin/insulin-11
3.1	74	27.2	58	2.3	AAK08499	Insulin/insulin-11
3.2	74	27.2	61	2.3	AAK08455	Insulin/insulin-11
3.3	74	27.2	61	2.3	AAK08500	Insulin/insulin-11
3.4	74	27.2	64	2.3	AAK08501	Insulin/insulin-11
3.5	74	27.2	65	2.3	AAK08456	Insulin/insulin-11
3.6	74	27.2	72	2.3	AAK08453	Insulin/insulin-11
3.7	74	27.2	72	2.3	AAK08498	Insulin/insulin-11
3.8	73.5	27.0	44	2.3	AAK08447	Insulin/insulin-11
3.9	73.5	27.0	79	2.3	AAK08448	Insulin/insulin-11
4.0	73.5	27.0	79	2.3	AAK088449	Insulin/insulin-11
4.1	73.5	27.0	84	2.3	AAK08450	Insulin/insulin-11
4.2	73	26.8	13	2.2	AAK019920	Bone Gla protein F
4.3	73	26.8	13	2.3	AAK18385	Human BGP pepide
4.4	72	26.5	198	2.2	AAK18386	Bovine gamma-cryst
4.5	72	26.5	198	2.2	AAK646849	Bovine gamma-cryst

## ALIGNMENTS

RESULT 1			
AA032936	AA032936 standard; peptide; 49 AA.		
XX			
AC	AA032936;		
XX			
DT	05-JUL-1993 (first entry)		
XX			
DE	21, 24, Glu human osteocalcin peptide.		
XX			
KW	Gamma-carboxyglutamic acid; OS.		
XX			
OS	Synthetic.		
XX			
PH	Key	Location/Qualifiers	
FT	Misc-difference 21	/label= OTHER	
FT		/note= "OTHER= gamma-carboxyglutamic acid"	
FT	Misc-difference 24	/label= OTHER	
FT		/note= "OTHER= gamma-carboxyglutamic acid"	
XX			
PN	<u>JR05032697-A.</u>		
XX			
PD	09-FEB-1993.		
XX			
PF	31-JUL-1991; 91JP-0213251.		
XX			
PR	31-JUL-1991; 91JP-0213251.		
XX			
PA	(TEIJU ) TEIJIN LTD.		
XX			
DR	WPI; 1993-088665/11.		

```

XX Synthetic human osteo:calcin for standard in determ. of natural
PT osteo:calcin - prepd. by introducing gamma-carboxy:glutamic acid
PT as fluorenyl protected gp.
XX
PS Claim 1, Page 2; 10pp; Japanese.
XX
CC The synthetic 21, 24, gamma-carboxyglutamic acid form of human
CC osteocalcin (OS) was produced by introducing protected Gla. The
CC substance may be produced in high yield and is useful as standard
CC for the determination of human OS.
CC See also AAR32937.
XX
SQ Sequence 49 AA;

Query Match 97.8%; Score 266; DB 14; Length 49;
Best Local Similarity 98.0%; Pred. No. 7.4e-30;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDCELDADHIGFQEAAYRRFGPV 49
   |||||||
Db 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDCELDADHIGFQEAAYRRFGPV 49

RESULT 2
AAR32937
ID AAR32937 standard; peptide; 49 AA.
XX
AC AAR32937;
XX
DT 05-JUL-1993 (first entry)
XX
DE 17, 21, 24, Gla human osteocalcin peptide.
XX
KM Gamma-carboxyglutamic acid; OS.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 17 /label= OTHER
FT /note= "OTHER= gamma-carboxyglutamic acid"
FT Misc-difference 21 /label= OTHER
FT /note= "OTHER= gamma-carboxyglutamic acid"
FT Misc-difference 24 /label= OTHER
FT /note= "OTHER= gamma-carboxyglutamic acid"
XX
PN JP05032627-A.
XX
PD 09-FEB-1993.
XX
PF 31-JUL-1991; 91JP-0213251.
XX
PR 31-JUL-1991; 91JP-0213251.
XX
PA (TEIJ ) TEIJIN LTD.
XX
DR WPI; 1993-088665/11.
XX
PT Synthetic human osteo:calcin for standard in determ. of natural
PT osteo:calcin - prepd. by introducing gamma-carboxy:glutamic acid
PT as fluorenyl protected gp.
XX
PS Claim 2; Page 2; 10pp; Japanese.
XX
CC The synthetic 17, 21, 24, gamma-carboxyglutamic acid form of human
CC osteocalcin (OS) was produced by introducing protected Gla. The
CC substance may be produced in high yield and is useful as standard
CC for the determination of human OS.
CC See also AAR32936.
XX

```

```

SQ Sequence 49 AA;

Query Match 97.8%; Score 266; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. No. 7.4e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDCELDADHIGFQEAAYRRFGPV 49
   |||||||
Db 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDCELDADHIGFQEAAYRRFGPV 49

RESULT 3
AAW34263
ID AAW34263 standard; peptide; 49 AA.
XX
AC AAW34263;
XX
DT 23-APR-1998 (first entry)
XX
DE Glu17-osteocalcin peptide #1.
XX
KM Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;
KM osteoporosis; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 21 /note= "gamma-carboxyglutamic acid"
XX
PN WO9738309-A1.
XX
PD 16-OCT-1997.
XX
PF 10-APR-1997; 97WO-JP01246.
XX
PR 27-FEB-1997; 97JP-0043331.
PR 10-APR-1996; 96JP-0088608.
XX
PA (EISA ) EISAI CO LTD.
XX
PI Kimura T, Morimoto S, Sakakibara S;
XX
DR WPI; 1997-512875/47.
XX
PT Antibody specific for Glu17-osteocalcin, or its fragment - for
PT diagnosis of bone disorders such as osteoporosis
XX
PS Claim 3; Page 16; 28pp; Japanese.
XX
CC This sequence represents a human Glu17-osteocalcin peptide. This sequence
CC is recognised by the antibody of the invention. The antibody of the
CC invention is an anti-Glu17-osteocalcin antibody or its fragment, which
CC binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The
CC antibody can be used for the diagnosis of bone related disorders, such as
CC osteoporosis.
XX
SQ Sequence 49 AA;

Query Match 97.8%; Score 266; DB 18; Length 49;
Best Local Similarity 95.9%; Pred. No. 7.4e-30;
Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDCELDADHIGFQEAAYRRFGPV 49
   |||||||
Db 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDCELDADHIGFQEAAYRRFGPV 49

RESULT 4
AAW34264
ID AAW34264 standard; peptide; 49 AA.
XX
AC AAW34264;
XX

```

XX 23-APR-1998 (first entry)  
XX  
XX  
DE Glu17-osteocalcin peptide #2.  
XX  
XX Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
XX osteoporosis; human.  
XX  
OS Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX Modified-site 21  
XX /note= "gamma-carboxyglutamic acid"  
XX Modified-site 24  
XX /note= "gamma-carboxyglutamic acid"  
XX  
XX W09738309-A1.  
XX  
XX 16-OCT-1997.  
XX  
XX 10-APR-1997; 97WO-JP01246.  
XX  
XX 27-FEB-1997; 97JP-0043331.  
XX 10-APR-1996; 96JP-0088608.  
XX  
XX (EISA ) EISAI CO LTD.  
XX  
XX Kimura T, Morimoto S, Sakakibara S;  
XX  
XX WPI; 1997-512875/47.  
XX  
XX  
XX Antibody specific for Glu17-osteocalcin, or its fragment - for  
XX diagnosis of bone disorders such as osteoporosis  
XX  
XX Claim 3; Page 16-17; 28pp; Japanese.  
XX  
XX This sequence represents a human Glu17-osteocalcin peptide. This sequence  
XX is recognised by the antibody of the invention. The antibody of the  
XX invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
XX binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The  
XX antibody can be used for the diagnosis of bone related disorders, such as  
XX osteoporosis.  
XX  
XX Sequence 49 AA:  
SQ  
Query Match 97.8%; Score 266; DB 18; Length 49;  
Best Local Similarity 98.0%; Pred. No. 7.4e-30;  
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 YLYQWLGAPVPPPLKRRXVCXINPDCDELADHIGFQEA YRRFGPV 49  
DB 1 YLYQWLGAPVPPPLKRRXVCXINPDCDELADHIGFQEA YRRFGPV 49  
RESULT 5  
AAW34265  
ID AAW34265 standard; peptide: 49 AA.  
XX  
XX AAW34265;  
XX  
XX 23-APR-1998 (first entry)  
XX  
XX Gla21-osteocalcin peptide #1.  
XX  
XX Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
XX osteoporosis; human.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 17  
XX /note= "gamma-carboxyglutamic acid"  
XX Modified-site 21

priority  
date 8/15/97  
page no 99d

FT /note= "gamma-carboxyglutamic acid"  
XX  
XX W09738309-A1.  
XX  
XX 16-OCT-1997.  
XX  
XX 10-APR-1997; 97WO-JP01246.  
XX  
XX 27-FEB-1997; 97JP-0043331.  
XX 10-APR-1996; 96JP-0088608.  
XX  
XX (EISA ) EISAI CO LTD.  
XX  
XX Kimura T, Morimoto S, Sakakibara S;  
XX  
XX WPI; 1997-512875/47.  
XX  
XX  
XX Antibody specific for Glu17-osteocalcin, or its fragment - for  
XX diagnosis of bone disorders such as osteoporosis  
XX  
XX Claim 4; Page 17; 28pp; Japanese.  
XX  
XX This sequence represents a human Gla21-osteocalcin peptide. This sequence  
XX is recognised by the antibody of the invention. The antibody of the  
XX invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
XX binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The  
XX antibody can be used for the diagnosis of bone related disorders, such as  
XX osteoporosis.  
XX  
XX Sequence 49 AA:  
SQ  
Query Match 97.8%; Score 266; DB 18; Length 49;  
Best Local Similarity 98.0%; Pred. No. 7.4e-30;  
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 YLYQWLGAPVPPPLKRRXVCXINPDCDELADHIGFQEA YRRFGPV 49  
DB 1 YLYQWLGAPVPPPLKRRXVCXINPDCDELADHIGFQEA YRRFGPV 49  
RESULT 6  
AAW34266  
ID AAW34266 standard; peptide: 49 AA.  
XX  
XX AAW34266;  
XX  
XX 23-APR-1998 (first entry)  
XX  
XX Gla21-osteocalcin peptide #2.  
XX  
XX Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
XX osteoporosis; human.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 17  
XX /note= "gamma-carboxyglutamic acid"  
XX Modified-site 21  
XX /note= "gamma-carboxyglutamic acid"  
XX Modified-site 24  
XX /note= "gamma-carboxyglutamic acid"  
XX  
XX W09738309-A1.  
XX  
XX 16-OCT-1997.  
XX  
XX 10-APR-1997; 97WO-JP01246.  
XX  
XX 27-FEB-1997; 97JP-0043331.  
XX 10-APR-1996; 96JP-0088608.  
XX  
XX (EISA ) EISAI CO LTD.  
XX

```
XX Kimura T, Morimoto S, Sakakibara S;
XX
XX WPI; 1997-512875/47.
XX
XX Antibody specific for Glu17-osteocalcin, or its fragment - for
XX diagnosis of bone disorders such as osteoporosis
XX
XX Claim 4; Page 18; 28pp; Japanese.
XX
XX This sequence represents a human Gla21-osteocalcin peptide. This sequence
XX is recognised by the antibody of the invention. The antibody of the
XX invention is an anti-Glu17-osteocalcin antibody or its fragment, which
XX binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The
XX antibody can be used for the diagnosis of bone related disorders, such as
XX osteoporosis.
XX
XX Sequence 49 AA;
SQ
Query Match 97.8%; Score 266; DB 18; Length 49;
Best Local Similarity 100.0%; Pred. No. 7.4e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDDELADHIGFQEAAYRRFGPV 49
DB 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDDELADHIGFQEAAYRRFGPV 49
RESULT 7
AAW99811
ID AAW99811 standard; Protein; 49 AA.
AC AAW99811;
XX
XX 08-JUN-1999 (first entry)
XX
XX Human osteocalcin fragment.
XX
XX Human; osteocalcin; urine; gamma-carboxylated; bone disorder; hOC;
XX bone turnover; puberty; menopause; detection.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 17
XX Modified-site /note= "possibly gamma-carboxylated"
XX Modified-site 21
XX Modified-site /note= "possibly gamma-carboxylated"
XX Modified-site 24
XX Modified-site /note= "possibly gamma-carboxylated"
XX
XX WO99058-A1.
XX
XX 25-FEB-1999.
XX
XX 24-JUN-1998; 98WO-FI00550.
XX
XX 15-AUG-1997; 97FI-0003371.
XX
XX (HELL/) HELLMAN J.
XX (KAER/) KAERKONEN S.
XX (KARE/) KARP M.
XX (LOEV/) LOEVGREN T.
XX (PET/) PETERSSON K.
XX (VAEA/) VAEANEN H K.
XX
XX Hellman J, Kaeronen S, Karp M, Loevgren T, Pettersson K;
XX Vaenenen HK;
XX
XX WPI; 1999-180972/15.
XX N-PSDB; AAX19767.
XX
XX Human osteocalcin peptide fragments - useful for detecting
```

```
PT metabolic bone turnover rates and disorders
XX
XX Claim 1; Fig 1A; 49pp; English.
XX
XX An assay has been developed for the detection of gamma-carboxylated
XX human osteocalcin found in urine, where the glutamic acid residues at
XX positions 17, 21 and 24 are gamma-carboxylated. The assay is used to
XX measure the rate of bone turnover (formation and/or resorption) and/or
XX for investigating metabolic bone disorders in individuals. Especially
XX mentioned, the assays can detect differences between adults and children
XX going through puberty, pre- and post-menopausal women and children
XX having high bone turnover. The assay allows more sensitive detection of
XX human osteocalcin (hOC) in urine than in serum. Significant increases in
XX hOC (40-48%) were detected in serum in menopause (normal increase is
XX 30-50% above pre-menopausal women). Detection of urine hOC using the
XX assay showed an increase as high as 75-79%. The present sequence
XX represents a fragment of hOC.
XX
XX Sequence 49 AA;
SQ
Query Match 97.8%; Score 266; DB 20; Length 49;
Best Local Similarity 93.9%; Pred. No. 7.4e-30;
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDDELADHIGFQEAAYRRFGPV 49
DB 1 YLYQWLGAAPVPPDPLXPRRXVCXLNPDDELADHIGFQEAAYRRFGPV 49
RESULT 8
AAR10146
ID AAR10146 standard; peptide; 51 AA.
XX
XX AAR10146;
XX
XX 27-MAR-1991 (first entry)
XX
XX Human osteocalcin precursor polypeptide (I).
XX
XX Human osteocalcin precursor polypeptide; OC; carboxypeptidase B;
XX calcium; vitamin K; bone formation; dysbolism.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 1..1 /note= "residue may be omitted"
XX Misc-difference 51..51 /note= "residue may be omitted"
XX
XX JP02201294-A.
XX
XX 03-DEC-1990.
XX
XX 22-AUG-1989; 89JP-0214239.
XX
XX 06-DEC-1988; 88JP-0306931.
XX 22-AUG-1989; 89JP-0214239.
XX
XX (TAKA-) TAKARA SYUZO KK.
XX
XX WPI; 1991-018865/03.
XX
XX Human osteocalcin (OC) precursor polypeptide - used to prepare
XX purified OC by digestion with carboxypeptidase B.
XX
XX Claim 1; Page 1; 11pp; Japanese.
XX
XX To K1 is attached H and to K51 is attached OH.
XX A novel gene encoding human OC precursor polymer was inserted into
XX plasmid pOC 980. E.Coli HB101 was transformed with this plasmid, and
XX made to efficiently produce the objective polymer. The polymer was
XX treated with lysylendopeptidase B to remove spacers. The obtained
```

CC human precursor was further treated with carboxy peptidase to prepare a  
CC purified human OC.  
CC OC is a calcium bonded protein (depending on vitamin K) produced in the  
CC bone. It is thought to be a promoting factor during the bone  
CC formation and may be used to treat diseases due to dyspolism of the  
CC bone.  
CC See also AAR10147 and AAQ10193-98.  
XX  
XX  
SQ Sequence 51 AA;  
Query Match 97.8%; Score 266; DB 12; Length 51;  
Best Local Similarity 93.9%; Pred. No. 7, 8e-30;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 YLQWLGAPVPPDPLKPRRXVCXKLNPCDELADHIGFOEAYRRFGPV 49  
DB 2 YLQWLGAPVPPDPLKPRRXVCXKLNPCDELADHIGFOEAYRRFGPV 50  
RESULT 9  
AAW76094  
ID AAW76094 standard; Protein; 98 AA.  
XX  
XX AAW76094;  
XX  
XX 21-DEC-1998 (first entry)  
DT  
XX  
DE Human osteocalcin protein.  
XX  
XX Promoter; tissue-specific gene expression; skeletal tissue;  
XX stem cell; bone; cartilage; osteocalcin; hOC gene; human;  
XX osteoporosis; osteopenia; osteosarcoma; cancer; metastasis;  
XX gene therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO9839427-AZ.  
PN  
XX  
XX 11-SEP-1998.  
PD  
XX  
XX 06-MAR-1998; 98MO-US04421.  
PF  
XX  
XX 06-MAR-1997; 97US-0039839.  
PR  
XX  
XX (UYMA-) UNITV MASSACHUSETTS.  
PA  
XX  
XX Frenkel B, Hou Z, Iian JB, Nilsson S, Peters S;  
PI Quesenberry P, Stein GS, Stein JL;  
PI  
XX  
XX WPI: 1998-495839/42.  
DR N-PSDB; AAV46429.  
XX  
XX  
XX Expression of exogenous genes in differentiated cells - by  
XX transducing pluripotent stem cells capable of maturing into  
XX differentiated cells with nucleic acid comprising exogenous gene,  
XX useful for, e.g. treatment of osteoporosis  
XX  
XX  
XX Disclosure; Page 33-34; 63pp; English.  
XX  
XX This is the amino acid sequence of the human osteocalcin, deduced  
XX from the coding exons of the human hOC gene (see AAV46429). The  
XX invention pertains to a method for expressing endogenous genes in  
XX differentiated cells of a specific type. The method involves  
XX contacting pluripotent stem cells capable of maturing into  
XX differentiated cells with a nucleic acid comprising an exogenous  
XX gene linked to a regulatory element capable of controlling expression  
XX of the exogenous gene in the differentiated cells. A population of  
XX transduced stem cells capable of maturing into differentiated cells  
XX expressing the exogenous gene is produced. Preferably, the  
XX differentiated cells are in a tissue of interest, such as bone or  
XX cartilage, and the exogenous gene is operably linked to at least one  
XX osteocalcin regulatory element derived from the hOC promoter. The  
XX exogenous gene can encode a therapeutic protein useful for treating

CC a disease, especially osteoporosis, osteopenia, osteosarcoma,  
CC primary malignancy or metastases (all claimed).  
XX  
XX  
SQ Sequence 98 AA;  
Query Match 97.8%; Score 266; DB 19; Length 98;  
Best Local Similarity 93.9%; Pred. No. 1, 6e-29;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 YLQWLGAPVPPDPLKPRRXVCXKLNPCDELADHIGFOEAYRRFGPV 49  
DB 50 YLQWLGAPVPPDPLKPRRXVCXKLNPCDELADHIGFOEAYRRFGPV 98  
RESULT 10  
AAR10147  
ID AAR10147 standard; Protein; 100 AA.  
XX  
XX AAR10147;  
XX  
XX 27-MAR-1991 (first entry)  
DT  
XX  
XX Human osteocarcin precursor polypeptide (II).  
DE  
XX  
XX Human osteocarcin precursor polypeptide; OC; carboxypeptidase B;  
XX calcium; vitamin K; bone formation; dysbolism.  
XX  
XX Homo sapiens.  
OS  
XX  
XX JP02201294-A.  
PN  
XX  
XX 03-DEC-1990.  
PD  
XX  
XX 22-AUG-1989; 89JP-0214239.  
PF  
XX  
XX 06-DEC-1988; 88JP-0306931.  
PR  
XX  
XX 22-AUG-1989; 89JP-0214239.  
XX  
XX (TAKA-) TAKARA SYUZO KK.  
XX  
XX WPI: 1991-018865/03.  
DR  
XX  
XX Human osteocarcin (OC) precursor polypeptide - used to prepare  
XX purified OC by digestion with carboxypeptidase B.  
PT  
XX  
XX Claim 5; Page 1; 11pp; Japanese.  
PS  
XX  
XX To K1 is attached H and to V100 is attached OH.  
XX A novel gene encoding human OC precursor polymer was inserted into  
XX plasmid pOC 980. E.coli HB101 was transformed with this plasmid, and  
XX made to efficiently produce the objective polymer. The polymer was  
XX treated with lysylendopeptidase B to remove spacers. The obtained  
XX human precursor was further treated with carboxy peptidase to prepare a  
XX purified human OC.  
XX OC is a calcium bonded protein (depending on vitamin K) produced in the  
XX bone. It is thought to be a promoting factor during the bone  
XX formation and may be used to treat diseases due to dyspolism of the  
XX bone.  
XX See also AAR10147 and AAQ10193-98.  
XX  
XX  
SQ Sequence 100 AA;  
Query Match 97.8%; Score 266; DB 12; Length 100;  
Best Local Similarity 93.9%; Pred. No. 1, 7e-29;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 YLQWLGAPVPPDPLKPRRXVCXKLNPCDELADHIGFOEAYRRFGPV 49  
DB 52 YLQWLGAPVPPDPLKPRRXVCXKLNPCDELADHIGFOEAYRRFGPV 100  
RESULT 11  
AAU10687

[illegible]

FT		/label= signal_peptide 23..44
FT	Peptide	/label= pro-peptide 45...48
FT	Cleavage-site	49..97
FT	Protein	/label= osteocalcin
XX		
PB	EPA63571-A.	
XX		
PD	02-JAN-1992.	
XX		
PF	20-JUN-1991;	91EP-0110173.
XX		
PR	30-NOV-1990; 90JP-0330146. 20-JUN-1990; 90JP-0159909.	
XX	(TEIJ ) TEIJJN KK.	
XX		
PI	Eguchi H, Kamimura TF, Sugiyama T, Hosoda K; WPT; 1992-009183/02.	
DR	N-PDB; AAQ20210.	
XX		
PT	Human osteocalcin produ. - using DNA coding for human osteocalcin fusion protein for expression in host cells	
PS	Clam 15; Fig 7; 53pp; English.	
CC	This sequence is a specific example of a claimed generic fusion protein comprising human osteocalcin. The pro-peptide is recognised by an enzyme capable of Glu to Gla conversion on human osteocalcin. The recombinant protein was obtained by culturing host cells transformed with a vector containing the synthetic coding sequence. CC The Glu residues could then be converted to Gla ( <i>i.e.</i> gamma-carboxyglutamic acid) and the osteocalcin sequence cleaved from the propeptide. The mature protein is suitable for use in immunossays and as a drug for treatment of bone metabolism disorders. See also AAR20044-6.	
SQ	Sequence     97 AA;	
Query Match	95.2%; Score 259; DB 13; Length 97; Best Local Similarity      91.8% Pred. No. 1,5e+28;	
Matches	45; Conservative     0; Mismatches     4; Indels     0; Gaps     0;	
OY	1 YLYQMIGAVPYPDPPLXRRXYCXINPCDDELADHIGFGAYRFFGVPV 49      iiiiiiiii             iiiiiiiii      49 YLYQMIGAVPYPPDDPLEPRREVCENLPCDDELDADHGIFGEAYRGFYGPV 97	
Db		
RESULT 13		.
AAM01681	ID	AAM01681 standard; protein; 49 AA.
AC	AAM01681;	
DT	01-APR-1997 (first entry)	
DE	Bone Gla protein.	
KM	BGP; bone gla protein; osteocalcin; Vitamin K-dependent protein; bone matrix; therapy; diagnosis; assay; metabolic bone disease.	
OS	Homo sapiens.	
FH	Key Location/Qualifiers	
FT	Misc-difference 1 /note= "underlined in specification"	
FT	Misc-difference 3 /note= "underlined in specification"	
FT	Misc-difference 12 /note= "underlined in specification"	
FT	Modified-site 21	



```

FT FT /label= OTHER
FT /note= "gamma-carboxyglutamic acid"
FT MISC-difference 23..29
FT /note= "underlined in specification"
FT Modified-site 24
FT /label= OTHER
FT MISC-difference 42
FT /note= "gamma-carboxyglutamic acid"
FT MISC-difference 46
FT /note= "underlined in specification"
FT /note= "underlined in specification"
PN DE4340597-A1.
PD 01-JUN-1995.
PE 29-NOV-1993; 93DE-4340597.
PR 29-NOV-1993; 93DE-4340597.
XX
XX (HENN-) HENNING BERLIN GMBH.
XX (BRAH-) BRAHMS DIAGNOSTICA GMBH.
XX Bergmann A, Weckermann R;
XX WPI; 1995-201516/27.
DR
XX Determn. of osteocalcin in serum or plasma - with addn. of divalent
XX metal ions to inhibit decompn. of the protein, useful in therapy
XX and diagnosis of bone disease
XX
XX PS Disclosure; Column 1; 9pp; German.
XX
XX CC A method for determining concentration of osteocalcin in serum or plasma
XX CC is improved with addition of divalent metal ions to inhibit decomposition
XX CC of the protein. The method is useful in therapy and diagnosis of bone
XX CC disease. The present sequence is osteocalcin (a vitamin K-dependent
XX CC protein that is a component of the bone matrix, or alternatively bone Gla
XX CC protein).
XX
XX SQ Sequence 49 AA;
XX
XX Query Match 94.9%; Score 258; DB 16; Length 49;
XX Best Local Similarity 95.9%; Pred. No. 9.8e-29;
XX Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 YLYQWLGAPVPPPLXPRRXVXCLNPDCELADHIGFOEAYRRFGPV 49
Db 1 YLYQWLGAPVPPPLXPRRXVXCLNPDCELADHIGFOEAYRRFGPV 49
RESULT 14
AAU10688
ID AAU10688 standard; Protein; 100 AA.
XX
XX AAU10688;
XX
XX 14-FEB-2002 (first entry)
XX
XX DE Polymorphic variant of human BGLAP protein.
XX
XX Human; single nucleotide polymorphism; SNP; BGLAP; chromosome 1q25-q31;
XX bone gamma carboxyglutamate protein; haplotyping; genotyping;
XX osteoporosis; osteopathic; variant.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX MISC-difference 82
XX /note= "Substitution of Glu to Lys"
XX MISC-difference 94
XX /note= "Substitution of Arg to Gln"
XX

```

```

PN WO200177131-A2.
XX
XX PD 18-OCT-2001.
XX
XX PF 11-APR-2001; 2001WO-US12010.
XX
XX PR 11-APR-2000; 2000US-195840P.
XX
XX (GENA-) GENMAISSANCE PHARM INC.
XX
XX Benlivena SC, Chew A, Choi JY, Koshy B, Rounds E, Stephens JC;
XX WPI; 2002-041288/05.
XX
XX PT New haplotypes of the human bone gamma carboxyglutamate protein gene,
XX useful to diagnose and treat diseases associated with the gene such as
XX osteoporosis
XX
XX PS Claim 27; Page -: 53pp; English.
XX
XX CC The present invention relates to novel single nucleotide polymorphisms
XX CC (SNPs) in the human bone gamma carboxyglutamate protein (BGLAP) gene
XX CC located on chromosome 1q25-q31, and methods for haplotyping and/or
XX CC genotyping the BGLAP gene in an individual. The methods of the
XX CC invention make use of allele-specific oligonucleotides (ASOs) as probes
XX CC and primers and/or primer-extension oligonucleotides for detecting the
XX CC BGLAP gene polymorphisms. The polymucleotides and screened compounds are
XX CC useful for (developing) treatment of diseases associated with BGLAP
XX CC activity, such as osteoporosis. The present sequence represents a
XX CC polymorphic variant of the BGLAP protein (AAU10687).
XX CC Note: The present sequence is not given in the specification but is
XX CC created by the indexer from the information given in the patent.
XX
XX SQ Sequence 100 AA;
XX
XX Query Match 94.9%; Score 258; DB 23; Length 100;
XX Best Local Similarity 89.8%; Pred. No. 2.2e-28;
XX Matches 44; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 YLYQWLGAPVPPPLXPRRXVXCLNPDCELADHIGFOEAYRRFGPV 49
Db 52 YLYQWLGAPVPPPLXPRRXVXCLNPDCELADHIGFOEAYRRFGPV 100
RESULT 15
AAB91917
ID AAB91917 standard; Peptide; 47 AA.
XX
XX AAB91917;
XX
XX 22-JUN-2001 (first entry)
XX
XX DE Bone Gla protein peptide SEQ ID NO:1093.
XX
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
XX blood component; modification; succinimidyl; maleimide group; amino;
XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
XX OS Homo sapiens.
XX
XX OS Synthetic.
XX
XX PN WO200069900-A2.
XX
XX PD 23-NOV-2000.
XX
XX PF 17-MAY-2000; 2000WO-US13576.
XX
XX PR 17-MAY-1999; 99US-0134406.
XX PR 10-SEP-1999; 99US-0153406.
XX PR 15-OCT-1999; 99US-0159783.
XX
XX (CONU-) CONUICHEM INC.
XX

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:39:45 : Search time 10 seconds  
(without alignments)  
79.587 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 49  
Sequence: 1 YLYQWLGAVPPDPLEPRR.....DELADHIGFEAVRRYGV 49

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 103943 seqs, 16242309 residues

Word size : 0

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	26.5	13	10	US-09-822-485-29
2	13	26.5	13	10	US-09-801-968-21
3	13	26.5	136	10	US-09-858-349-2
4	6	12.2	61	9	US-10-001-876-146
5	6	12.2	70	9	US-09-895-913A-364
6	6	12.2	70	10	US-09-815-242-11343
7	6	12.2	103	10	US-09-841-133-501
8	6	12.2	116	10	US-09-775-932-16
9	6	12.2	139	10	US-09-969-834-4
10	6	12.2	934	10	US-09-788-657-19
11	6	12.2	1690	10	US-09-788-043C-5
12	5	10.2	9	10	US-09-924-400-138
13	5	10.2	9	10	US-09-810-936-138
14	5	10.2	9	10	US-09-429-755-138
15	5	10.2	23	9	US-09-924-400-133
16	5	10.2	23	10	US-09-864-761-34069
17	5	10.2	23	10	US-09-810-936-133
18	5	10.2	23	10	US-09-429-755-133
19	5	10.2	24	10	US-09-922-261-152

20	5	10.2	26	10	US-09-821-984-38	Sequence 38, Appl
21	5	10.2	28	10	US-09-799-983-17	Sequence 17, Appl
22	5	10.2	38	10	US-09-864-761-38378	Sequence 38378, A
23	5	10.2	40	10	US-09-925-299-1355	Sequence 1355, A
24	5	10.2	46	10	US-09-888-615-117	Sequence 117, Appl
25	5	10.2	47	10	US-09-864-761-33570	Sequence 33570, A
26	5	10.2	47	10	US-09-864-761-37719	Sequence 37719, A
27	5	10.2	53	10	US-09-873-880-12	Sequence 12, Appl
28	5	10.2	57	10	US-09-864-761-36551	Sequence 36551, A
29	5	10.2	60	10	US-09-925-297-514	Sequence 514, Appl
30	5	10.2	63	10	US-09-925-550-1110	Sequence 1110, Ap
31	5	10.2	64	10	US-09-764-860-451	Sequence 451, Appl
32	5	10.2	65	10	US-09-925-297-742	Sequence 742, Appl
33	5	10.2	65	10	US-09-764-847-647	Sequence 647, Appl
34	5	10.2	67	10	US-09-864-761-45197	Sequence 45197, A
35	5	10.2	69	10	US-09-804-156-20	Sequence 4, Appl
36	5	10.2	69	10	US-09-867-550-1321	Sequence 37, Appl
37	5	10.2	69	10	US-09-764-847-888	Sequence 888, Appl
38	5	10.2	74	10	US-09-925-300-509	Sequence 509, Appl
39	5	10.2	84	10	US-09-764-847-592	Sequence 592, Appl
40	5	10.2	86	10	US-09-864-761-33499	Sequence 33499, A
41	5	10.2	88	10	US-09-925-300-1321	Sequence 1321, Ap
42	5	10.2	89	10	US-09-764-860-437	Sequence 437, Appl
43	5	10.2	92	10	US-09-925-300-1795	Sequence 1795, Ap
44	5	10.2	94	10	US-09-764-877-1490	Sequence 1490, Ap
45	5	10.2	97	10	US-09-873-880-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-09-822-485-29  
; Sequence 29, Application US/09822485  
; Patent No. US20020001825A1  
; GENERAL INFORMATION:  
; APPLICANT: Itch, No. US20020001825A1yuyuk1  
; TITLE OF INVENTION: No. US20020001825A1el Fibroblast Growth Factor-Like Polypeptid  
; FILE REFERENCE: 08035.0001-01000  
; CURRENT APPLICATION NUMBER: US/09/822,485  
; CURRENT FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:protein tag  
US-09-822-485-29

Query Match 26.5%; Score 13; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 8.3e-08;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDLEPR 19  
Db 1 GAVPYPPDLEPR 13

RESULT 2  
US-09-801-968-21  
; Sequence 21, Application US/09801968  
; Patent No. US20020082205A1  
; GENERAL INFORMATION:  
; APPLICANT: Itch, No. US20020082205A1yuyuk1  
; APPLICANT: Kavanaugh, W. Michael  
; TITLE OF INVENTION: HUMAN FGF-23 GENE AND GENE EXPRESSION  
; PRODUCTS  
; FILE REFERENCE: PP-17150.001/201130.40901  
; CURRENT APPLICATION NUMBER: US/09/801,968  
; CURRENT FILING DATE: 2001-03-07  
; NUMBER OF SEQ ID NOS: 46

SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: E tag  
US-09-801-968-21

Query Match  
Best Local Similarity 100.0%; Score 13; DB 10; Length 13;  
Pred. No. 8.3e-08;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDLEPR 19  
Db 1 GAVPYPPDLEPR 13

RESULT 3  
US-09-858-349-2  
; Sequence 2, Application US/09858349  
; Patent No. US20020012909A1  
; GENERAL INFORMATION:  
; APPLICANT: PLAKSIN, Daniel  
; TITLE OF INVENTION: SMALL FUNCTIONAL UNITS OF ANTIBODY HEAVY CHAIN VARIABLE REGIONS  
; FILE REFERENCE: 87534-2800  
; CURRENT APPLICATION NUMBER: US/09/858,349  
; CURRENT FILING DATE: 2000-06-02  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: mouse hybridoma specific for H-2D + RGPGRFVTL peptide  
; FEATURE:  
; NAME/KEY: MSC\_FEATURE  
; LOCATION: (99)..(107)  
; OTHER INFORMATION: variable  
US-09-858-349-2

Query Match  
Best Local Similarity 100.0%; Score 13; DB 10; Length 136;  
Pred. No. 6.7e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDLEPR 19  
Db 122 GAVPYPPDLEPR 134

RESULT 4  
US-10-001-876-146  
; Sequence 146, Application US/10001876  
; Patent No. US20020177140A1  
; GENERAL INFORMATION:  
; APPLICANT: Salceda, Susana  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Hevve  
; APPLICANT: Cafferey, Robert  
; APPLICANT: Ali, Shujath  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes in  
; FILE REFERENCE: DEX-0285  
; CURRENT APPLICATION NUMBER: US/10/001,876  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/252,186  
; PRIOR FILING DATE: 2000-11-21  
; NUMBER OF SEQ ID NOS: 211  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 146  
; LENGTH: 61  
; TYPE: PRT  
; ORGANISM: Homo sapien

US-10-001-876-146

Query Match  
Best Local Similarity 100.0%; Score 6; DB 9; Length 61;  
Pred. No. 6.9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LEPRRE 21  
Db 29 LEPRRE 34

RESULT 5  
US-09-895-913A-364  
; Sequence 364, Application US/09895913A  
; Patent No. US20020160456A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleantous, Harold  
; APPLICANT: Al-Garawi, Amal  
; APPLICANT: Miller, Charles  
; APPLICANT: Tomb, Jean Francois  
; APPLICANT: Comen, Raymond P.  
; TITLE OF INVENTION: Identification of polynucleotides  
; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in t  
; FILE REFERENCE: 06132/043002  
; CURRENT APPLICATION NUMBER: US/09/895,913A  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 08/881,227  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 364  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
US-09-895-913A-364

Query Match  
Best Local Similarity 100.0%; Score 6; DB 9; Length 70;  
Pred. No. 7.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EAYRRF 45  
Db 14 EAYRRF 19

RESULT 6  
US-09-815-242-11343  
; Sequence 11343, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 11343  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Helicobacter pylori  
US-09-815-242-11343

Query Match 12.2%; Score 6; DB 10; Length 70;  
Best Local Similarity 100.0%; Pred. No. 7.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EAYRRF 45  
|||||  
DB 14 EAYRRF 19

RESULT 7  
US-09-841-132-501  
Sequence 501, Application US/09841132  
Patent No. US20020061848A1  
GENERAL INFORMATION:  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Probst, Peter  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
FILE REFERENCE: 210121.469C8  
CURRENT APPLICATION NUMBER: US/09/841,132  
CURRENT FILING DATE: 2001-04-23  
NUMBER OF SEQ ID NOS: 599  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO: 501  
LENGTH: 103  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-09-841-132-501

Query Match 12.2%; Score 6; DB 10; Length 103;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLXQWL 6  
|||||  
DB 20 YLXQWL 25

RESULT 8  
US-09-775-932-16  
Sequence 16, Application US/09775932  
Patent No. US20020137671A1  
GENERAL INFORMATION:  
APPLICANT: University of British Columbia  
TITLE OF INVENTION: Production and use of Modified Cystatins  
FILE REFERENCE: 58069  
CURRENT APPLICATION NUMBER: US/09/775,932  
CURRENT FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: CA99/00717  
PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: 60/095,503  
PRIOR FILING DATE: 1998-08-05  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO: 16  
LENGTH: 116  
TYPE: PRT  
ORGANISM: Gallus sp.  
US-09-775-932-16

Query Match 12.2%; Score 6; DB 10; Length 116;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
|||||  
DB 8 LGAPVP 13

RESULT 9  
US-09-969-834-4  
Sequence 4, Application US/09969834  
Patent No. US20020102711A1  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE  
PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/969,834  
FILING DATE: 01-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/471,765  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/08/791,522  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 09/471,765  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0193 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 118195  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-969-834-4

Query Match 12.2%; Score 6; DB 10; Length 139;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
|||||  
DB 31 LGAPVP 36

RESULT 10  
US-09-788-657-19

```
; Sequence 19, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sasse, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 934
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-788-657-19
```

```
Query Match 12.2%; Score 6; DB 10; Length 934;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 7 GAPVY 12
Db 614 GAPVY 619
```

```
RESULT 11
US-09-788-043C-5
; Sequence 5, Application US/09788043C
; Patent No. US20020107361A1
; GENERAL INFORMATION:
; APPLICANT: Heller, Renu
; APPLICANT: Zhu, Fengrong
; APPLICANT: Klonowski, Paul
; TITLE OF INVENTION: No. US20020107361A1 Metalloproteases Having
; TITLE OF INVENTION: Thrombospondin Domains and Nucleic Acid Compositions
; FILE REFERENCE: ROCH-004
; CURRENT APPLICATION NUMBER: US/09/788,043C
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/184,152
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1690
; TYPE: PRT
; ORGANISM: human
US-09-788-043C-5
```

```
Query Match 12.2%; Score 6; DB 10; Length 1690;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 16 LEPRRE 21
Db 216 LEPRRE 221
```

```
RESULT 12
US-09-924-400-138
; Sequence 138, Application US/09924400
; Patent No. US20020165371A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
```

```
; APPLICANT: Misher, Lynda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.419C12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted HLA A2.1 Motifs (T-cell epitopes)
US-09-924-400-138
```

```
Query Match 10.2%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 39 QEAYR 43
Db 3 QEAYR 7
```

```
RESULT 13
US-09-810-936-138
; Sequence 138, Application US/09810936
; Patent No. US20020068285A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.419C11
; CURRENT APPLICATION NUMBER: US/09/810,936
; CURRENT FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 138
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted HLA A2.1 Motifs (T-cell epitopes)
US-09-810-936-138
```

```
Query Match 10.2%; Score 5; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.6e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 39 QEAYR 43
Db 3 QEAYR 7
```

```
RESULT 14
```

```
US-09-429-755-138
; Sequence 138, Application US/09429755A
; Patent No. US2002011467A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.419C6
; CURRENT APPLICATION NUMBER: US/09/429,755A
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 138
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted HLA A2.1 Motifs (T-cell epitopes)
US-09-429-755-138

Query Match          10.2%: Score 5; DB 10; Length 9;
Best Local Similarity 100.0%; Pred.No. 8.6e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      39 QEAYR 43
      |||||
Db      3 QEAYR 7

RESULT 15
US-09-924-400-133
; Sequence 133, Application US/09924400
; Patent No. US20020165371A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Lynda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aljun
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; APPLICANT: Deng, Ya
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.419C12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted Th Motifs (B-cell epitopes)
US-09-924-400-133

Query Match          10.2%: Score 5; DB 9; Length 23;
Best Local Similarity 100.0%; Pred.No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      39 QEAYR 43
      |||||
Db      10 QEAYR 14
```

```
RESULT 16
US-09-864-761-34069
; Sequence 34069, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmlca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34069
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO l29074.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
US-09-864-761-34069

Query Match          10.2%: Score 5; DB 10; Length 23;
```

Best Local Similarity 100.0%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 EVCEL 25  
      |||||  
DB 18 EVCEL 22

## RESULT 17

US-09-810-936-133  
; Sequence 133, Application US/09810936  
; Patent No. US2002006285A1  
; GENERAL INFORMATION:  
; APPLICANT: Fridakis, Tony N.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Smith, John M.  
; APPLICANT: Misher, Linda E.  
; APPLICANT: Dillon, David C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Day, Craig H.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C11  
; CURRENT APPLICATION NUMBER: US/09/810,936  
; CURRENT FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 334  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 133  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Predicted 7h Motifs (B-cell epitopes)  
US-09-810-936-133

Query Match 10.2%; Score 5; DB 10; Length 23;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
      |||||  
DB 10 QEAYR 14

## RESULT 18

US-09-429-755-133  
; Sequence 133, Application US/09429755A  
; Patent No. US2002011467A1  
; GENERAL INFORMATION:  
; APPLICANT: Fridakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Misher, Lynda  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Dillon, David C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C6  
; CURRENT APPLICATION NUMBER: US/09/429,755A  
; CURRENT FILING DATE: 1999-10-28  
; NUMBER OF SEQ ID NOS: 315  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 133  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Predicted 7h Motifs (B-cell epitopes)  
US-09-429-755-133

Query Match 10.2%; Score 5; DB 10; Length 23;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
      |||||  
DB 10 QEAYR 14

## RESULT 19

US-09-922-261-152  
; Sequence 152, Application US/09922261  
; Patent No. US2002011471A1  
; GENERAL INFORMATION:  
; APPLICANT: COGENE NEUROSCIENCE, Inc.  
; APPLICANT: Lo, Donald C.  
; APPLICANT: Barney, Shawn  
; APPLICANT: Thomas, Mary Beth  
; APPLICANT: Portbury, Stuart D.  
; APPLICANT: Puranam, Kasturi  
; APPLICANT: Katz, Lawrence C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING  
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING  
; TITLE OF INVENTION: CELL DEATH  
; FILE REFERENCE: 10001-005-999  
; CURRENT APPLICATION NUMBER: US/09/922,261  
; CURRENT FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: US/09/461,697  
; PRIOR FILING DATE: 1999-12-14  
; NUMBER OF SEQ ID NOS: 466  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 152  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-922-261-152

Query Match 10.2%; Score 5; DB 10; Length 24;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LEPRR 20  
      |||||  
DB 16 LEPRR 20

## RESULT 20

US-09-821-984-38  
; Sequence 38, Application US/09821984  
; Patent No. US2002004205A1  
; GENERAL INFORMATION:  
; APPLICANT: Consler, Thomas G.  
; APPLICANT: Iannone, Marie A.  
; APPLICANT: Gray, John G.  
; APPLICANT: Stimmel, Julia E.  
; TITLE OF INVENTION: METHOD OF INVESTIGATING FUNCTIONAL  
; TITLE OF INVENTION: MOLECULAR INTERACTIONS AND REAGENTS FOR USE THEREIN  
; FILE REFERENCE: 07083.0007U2  
; CURRENT APPLICATION NUMBER: US/09/821,984  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 60/193,826  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 38  
; LENGTH: 26  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/note -  
US-09-821-984-38



Query Match 10.2%; Score 5; DB 10; Length 26;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LADHI 36  
|||||  
DB 12 LADHI 16

RESULT 21  
US-09-799-983-17  
; Sequence 17, Application US/09799983  
; Patent No. US20010029033A1  
; GENERAL INFORMATION:  
; APPLICANT: Shamli, Paul  
; APPLICANT: Parker, Charles  
; TITLE OF INVENTION: NOVEL GENE RNO UPREGULATED IN LEUKEMIA CELLS BY NITRIC OXIDE AN  
; FILE REFERENCE: 1321.2.51  
; CURRENT APPLICATION NUMBER: US/09/799,983  
; CURRENT FILING DATE: 2001-03-06  
; PRIOR APPLICATION NUMBER: US 60/186,971  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-799-983-17

Query Match 10.2%; Score 5; DB 10; Length 28;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELA 33  
|||||  
DB 17 CDELA 21

RESULT 22  
US-09-864-761-38378  
; Sequence 38378, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomlca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30

0  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 38378  
; LENGTH: 38  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005630.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.2  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.5  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.9  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.7  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.7  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.3  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.5  
US-09-864-761-38378

Query Match 10.2%; Score 5; DB 10; Length 38;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GADVP 11  
|||||  
DB 24 GADVP 28

RESULT 23  
US-09-925-299-1355  
; Sequence 1355, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1355  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (25)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (34)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (36)

```
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1355
Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVP 11
    |||||
DB 28 GAVP 32

RESULT 24
US-09-888-615-117
; Sequence 117, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOMMAN, GREGORY
; APPLICANT: WHITE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-117

Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 VCELN 26
    |||||
DB 8 VCELN 12

RESULT 25
US-09-864-761-33570
; Sequence 33570, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33570
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP00111.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.82
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: AA757099.1, EVALU 6.00e-06
; OTHER INFORMATION: SWISSPROT HIT: Q10411, EVALU 2.70e+00
US-09-864-761-33570

Query Match
Best Local Similarity 10.2%; Score 5; DB 10; Length 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EPRRE 21
    |||||
DB 23 EPRRE 27

RESULT 26
US-09-864-761-37719
; Sequence 37719, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
```

;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: GB 24263.6  
;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 37719  
;; LENGTH: 47  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AP000043.1  
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.63  
;; OTHER INFORMATION: EST\_HUMAN HIT: AA57099.1, EVALU 6.00e-06  
;; OTHER INFORMATION: SWISSPROT HIT: Q10411, EVALU 2.70e+00  
US-09-864-761-37719  
Query Match 10.2%; Score 5; DB 10; Length 47;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 EPRE 21  
|||||  
Db 23 EPRE 27

RESULT 27  
US-09-873-880-12  
;; Sequence 12, Application US/09873880  
;; Patent No. US20020123118A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Sewall, Vincent  
;; APPLICANT: Falco, S. Carl  
;; APPLICANT: Allen, Stephen M.  
;; TITLE OF INVENTION: GLYCINE METABOLISM ENZYMES  
;; FILE REFERENCE: BB1192 US CIP  
;; CURRENT APPLICATION NUMBER: US/09/873,880  
;; PRIOR FILING DATE: 2001-06-04  
;; PRIOR APPLICATION NUMBER: 09/363,321  
;; PRIOR FILING DATE: JULY 28, 1999  
;; PRIOR APPLICATION NUMBER: 60/094,839  
;; PRIOR FILING DATE: JULY 31, 1998  
;; NUMBER OF SEQ ID NOS: 42  
;; SOFTWARE: Microsoft Office 97  
;; SEQ ID NO 12

;; LENGTH: 53  
;; TYPE: PRT  
;; ORGANISM: Zea mays  
US-09-873-880-12  
Query Match 10.2%; Score 5; DB 10; Length 53;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 DELAD 34  
|||||  
Db 46 DELAD 50

RESULT 28  
US-09-864-761-36551  
;; Sequence 36551, Application US/09864761  
;; Patent No. US20020048763A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Penn, Sharon G.  
;; APPLICANT: Rank, David R.  
;; APPLICANT: Hanzel, David K.  
;; APPLICANT: Chen, Wensheng  
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
;; FILE REFERENCE: Accmca-X-1  
;; CURRENT APPLICATION NUMBER: US/09/864,761  
;; PRIOR FILING DATE: 2001-05-23  
;; PRIOR APPLICATION NUMBER: US 60/180,312  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: US 60/207,456  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: US 09/632,366  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: GB 24263.6  
;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 36551  
;; LENGTH: 57  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AP000247.1  
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.67

```

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.74
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PER474, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.65
; OTHER INFORMATION: EST_HUMAN HIT: BE968985.1, EVALUATE 7.20e-01
; OTHER INFORMATION: SWISSPROT HIT: Q13009, EVALUATE 1.00e-31
US-09-864-761-36551

```

```

Query Match          10.2%; Score 5; DB 10; Length 57;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 26 NPDCD 30
    |||||
Db 14 NPDCD 18

```

```

RESULT 29
US-09-925-297-514
; Sequence 514, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 514
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-514

```

```

Query Match          10.2%; Score 5; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 45 FYGPV 49
    |||||
Db 35 FYGPV 39

```

```

RESULT 30
US-09-867-550-1110
; Sequence 1110, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad.

```

```

; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1110
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1110

```

```

Query Match          10.2%; Score 5; DB 10; Length 63;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 12 YPDPL 16
    |||||
Db 29 YPDPL 33

```

```

RESULT 31
US-09-764-860-451
; Sequence 451, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 451
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-860-451

```

```

Query Match          10.2%; Score 5; DB 10; Length 64;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 YOWLG 7
    |||||
Db 51 YOWLG 55

```

```

RESULT 32
US-09-925-297-742
; Sequence 742, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989

```

;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12  
;; NUMBER OF SEQ ID NOS: 928  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 742  
;; LENGTH: 65  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-925-297-742

Query Match  
Best Local Similarity 100.0%; Score 5; DB 10; Length 65;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 CELNP 27  
|||||  
DB 51 CELNP 55

RESULT 33  
US-09-764-847-647  
;; Sequence 647, Application US/09764847  
;; Patent No. US20020133767A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
;; FILE REFERENCE: PC009  
;; CURRENT APPLICATION NUMBER: US/09/764,847  
;; CURRENT FILING DATE: 2001-01-17  
;; Prior application data removed - consult PALM or file wrapper  
;; NUMBER OF SEQ ID NOS: 2003  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 647  
;; LENGTH: 65  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-764-847-647

Query Match  
Best Local Similarity 100.0%; Score 5; DB 10; Length 65;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 CELNP 27  
|||||  
DB 51 CELNP 55

RESULT 34  
US-09-864-761-45197  
;; Sequence 45197, Application US/09864761  
;; Patent No. US20020048763A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Penn, Sharon G.  
;; APPLICANT: Rank, David R.  
;; APPLICANT: Hanzel, David K.  
;; APPLICANT: Chen, Wensheng  
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
;; FILE REFERENCE: Aegonica-X-1  
;; CURRENT APPLICATION NUMBER: US/09/864,761  
;; CURRENT FILING DATE: 2001-05-23  
;; PRIOR APPLICATION NUMBER: US 60/180,312  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: US 60/207,456  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: US 09/632,366  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: GB 24263.6  
;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666

;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 45197  
;; LENGTH: 67  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

FEATURE:  
;; OTHER INFORMATION: MAP TO AL121673.27  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.83  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.86  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.83  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.85  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.81  
;; OTHER INFORMATION: SWISSPROT HIT: P26807, EVALUATE 2.00e+00  
;; OTHER INFORMATION: EST\_HUMAN HIT: AW438701.1, EVALUATE 1.90e+00  
US-09-864-761-45197

Query Match  
Best Local Similarity 100.0%; Score 5; DB 10; Length 67;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVP 11  
|||||  
DB 31 GAPVP 35

RESULT 35  
US-09-804-156-20  
;; Sequence 20, Application US/09804156  
;; Patent No. US20020068320A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Shi et al.  
;; TITLE OF INVENTION: Serine protease Polynucleotides, Polypeptides, and Antibodies  
;; FILE REFERENCE: PT005P4  
;; CURRENT APPLICATION NUMBER: US/09/804,156  
;; CURRENT FILING DATE: 2001-03-13  
;; PRIOR APPLICATION NUMBER: 60/189,025  
;; PRIOR FILING DATE: 2000-03-14  
;; NUMBER OF SEQ ID NOS: 38  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 20  
;; LENGTH: 69  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-804-156-20

Query Match  
Best Local Similarity 100.0%; Score 5; DB 10; Length 69;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 VCELN 26  
|||||  
DB 14 VCELN 18

RESULT 36  
US-09-804-156-37

; Sequence 37, Application US/09804156  
; Patent No. US20020068320A1

; GENERAL INFORMATION:

; APPLICANT: Shi et al.

; TITLE OF INVENTION: Serine protease Polynucleotides, Polypeptides, and Antibodies

; FILE REFERENCE: PTO05P4

; CURRENT APPLICATION NUMBER: US/09/804,156

; CURRENT FILING DATE: 2001-03-13

; PRIOR APPLICATION NUMBER: 60/189,025

; PRIOR FILING DATE: 2000-03-14

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 37

; LENGTH: 69

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-804-156-37

Query Match 10.2%; Score 5; DB 10; Length 69;

Best Local Similarity 100.0%; Pred. No. 86;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 VCELN 26  
|||||  
DB 14 VCELN 18

RESULT 37  
US-09-764-847-888

; Sequence 888, Application US/09764847

; Patent No. US20020132767A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC009

; CURRENT APPLICATION NUMBER: US/09/764,847

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2003

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 888

; LENGTH: 69

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-847-888

Query Match 10.2%; Score 5; DB 10; Length 69;

Best Local Similarity 100.0%; Pred. No. 86;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WIGAP 9  
|||||  
DB 47 WIGAP 51

RESULT 38  
US-09-925-302-509

; Sequence 509, Application US/09925302

; Patent No. US2002004941A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA104

; CURRENT APPLICATION NUMBER: US/09/925,302

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05918

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 896

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 509

; LENGTH: 74

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (31)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-302-509

Query Match 10.2%; Score 5; DB 10; Length 74;

Best Local Similarity 100.0%; Pred. No. 91;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 HIGFQ 39  
|||||  
DB 4 HIGFQ 8

RESULT 39  
US-09-764-847-592

; Sequence 592, Application US/09764847

; Patent No. US20020132767A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC009

; CURRENT APPLICATION NUMBER: US/09/764,847

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2003

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 592

; LENGTH: 84

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (41)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (81)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-847-592

Query Match 10.2%; Score 5; DB 10; Length 84;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPV 10  
|||||  
DB 65 LGAPV 69

RESULT 40  
US-09-864-761-33499

; Sequence 33499, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aeomica-x-1

```

; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33499
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035603.11
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 32
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 27
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 27
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.5
; OTHER INFORMATION: SWISSPROT HIT: P79103, EVALU 9.00e-30
; OTHER INFORMATION: EST_HUMAN HIT: W00563.1, EVALU 5.00e-34
; US-09-864-761-33499

Query Match          10.2%; Score 5; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 YPDP 16
    |||||
DB 53 YPDP 57

RESULT 41
; US-09-925-300-1321
; Sequence 1321, Application US/09925300
; Patent No. US20020151681A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1321
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-300-1321

Query Match          10.2%; Score 5; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 PREV 22
    |||||
DB 12 PREV 16

RESULT 42
; US-09-764-860-437
; Sequence 437, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 437
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-860-437

Query Match          10.2%; Score 5; DB 10; Length 89;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LEPR 20
    |||||
DB 30 LEPR 34

RESULT 43
; US-09-925-300-1795
; Sequence 1795, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentln Ver. 2.0
```

; SEQ ID NO 1795  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-300-1795

Query Match  
Best Local Similarity 10.2%; Score 5; DB 10; Length 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 YPDL 16  
Db 53 YPDL 57

RESULT 44  
US-09-764-877-1490  
; Sequence 1490, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1490  
; LENGTH: 94  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (63)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (94)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-877-1490

Query Match  
Best Local Similarity 10.2%; Score 5; DB 10; Length 94;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADH 35  
Db 68 ELADH 72

RESULT 45  
US-09-873-880-14  
; Sequence 14, Application US/09873880  
; Patent No. US20020123118A1  
; GENERAL INFORMATION:  
; APPLICANT: Sewalt, Vincent  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: GLYCINE METABOLISM ENZYMES  
; FILE REFERENCE: BBI192 US CIP  
; CURRENT APPLICATION NUMBER: US/09/873,880  
; CURRENT FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 09/363,321  
; PRIOR FILING DATE: July 28, 1999  
; PRIOR APPLICATION NUMBER: 60/094,839  
; PRIOR FILING DATE: July 31, 1998  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 14  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-873-880-14

Query Match  
Best Local Similarity 10.2%; Score 5; DB 10; Length 97;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 DELAD 34  
Db 46 DELAD 50

Search completed: December 4, 2002, 15:43:32  
Job time : 11 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:38:05 ; Search time 15 Seconds  
(without alignments)  
96.115 Million cell updates/sec

Title: US-09-462-931-2

Sequence: 1 YLYQWLGAAPVPPDPLERRR.....DELADHIGFOEAYRRFYGPV 49

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49	100.0	49	6	5434245-1	Patent No. 5434245
2	49	100.0	50	6	5434245-2	Patent No. 5434245
3	49	100.0	50	6	5434245-3	Patent No. 5434245
4	49	100.0	51	6	5434245-4	Patent No. 5434245
5	25	51.0	49	6	5164483-1	Patent No. 5164483
6	25	51.0	49	6	5164483-2	Patent No. 5164483
7	25	51.0	98	6	5164483-3	Patent No. 5164483
8	19	38.8	48	6	5168041-1	Patent No. 5168041
9	13	26.5	13	2	US-08-796-598-13	Sequence 13, Appl
10	13	26.5	13	2	US-08-447-175A-13	Sequence 13, Appl
11	13	26.5	13	2	US-08-943-915-6	Sequence 6, Appl
12	13	26.5	13	2	US-08-881-037-112	Sequence 112, Appl
13	13	26.5	13	3	US-08-652-816A-43	Sequence 43, Appl
14	13	14.3	13	4	US-09-142-974B-5	Sequence 5, Appl
15	13	14.3	13	4	US-08-466-120-2	Sequence 2, Appl
16	7	14.3	433	5	PCT-US94-07266-2	Sequence 2, Appl
17	7	14.3	440	1	US-08-333-338-8	Sequence 8, Appl
18	7	14.3	440	1	US-08-463-694-8	Sequence 8, Appl
19	7	14.3	440	1	US-08-694-501-8	Sequence 8, Appl
20	7	14.3	447	1	US-08-373-935-1	Sequence 1, Appl
21	6	12.2	12	3	US-08-968-747-7	Sequence 7, Appl
22	6	12.2	12	3	US-08-493-071-28	Sequence 28, Appl
23	6	12.2	139	2	US-08-791-522-4	Sequence 4, Appl
24	6	12.2	139	3	US-09-314-777-4	Sequence 4, Appl
25	6	12.2	934	1	US-08-457-176-2	Sequence 2, Appl
26	6	12.2	934	1	US-08-457-175-2	Sequence 2, Appl
27	6	12.2	934	3	US-08-709-784-1	Sequence 1, Appl

28	6	12.2	934	4	US-09-651-656-3	Sequence 3, Appl
29	6	12.2	934	4	US-09-650-855-3	Sequence 3, Appl
30	6	12.2	1050	3	US-09-045-632-49	Sequence 49, Appl
31	6	12.2	1050	3	US-09-045-632-50	Sequence 50, Appl
32	6	12.2	3170	2	US-07-642-734C-5	Sequence 5, Appl
33	6	12.2	3170	3	US-08-439-009A-5	Sequence 138, App
34	5	10.2	9	4	US-08-991-789A-138	Sequence 138, App
35	5	10.2	9	4	US-09-062-451-138	Sequence 138, App
36	5	10.2	9	4	US-09-598-326-138	Sequence 138, App
37	5	10.2	12	1	US-08-247-809A-10	Sequence 10, Appl
38	5	10.2	12	2	US-08-711-728-10	Sequence 10, Appl
39	5	10.2	13	5	PCT-US93-05647-18	Sequence 18, Appl
40	5	10.2	18	1	US-08-101-041A-4	Sequence 4, Appl
41	5	10.2	18	2	US-09-025-706-6	Sequence 6, Appl
42	5	10.2	18	4	US-09-025-622-6	Sequence 125, App
43	5	10.2	20	4	US-08-602-999A-125	Sequence 125, App
44	5	10.2	20	4	US-09-500-124-125	Sequence 12, Appl
45	5	10.2	22	1	US-08-403-378B-12	Sequence 12, Appl

## ALIGNMENTS

RESULT 1  
5434245-1  
;PATENT NO. 5434245  
; APPLICANT: KOYAMA, NOBUTO;KIMIZUKA, FUSAO;KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/993,980  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989  
; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989  
; SEQ ID NO:1:  
; LENGTH: 49  
5434245-1

Query Match 100.0%; Score 49; DB 6; Length 49;  
Best Local Similarity 100.0%; Pred. No. 5.8e-45;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 YLYQWLGAAPVPPDPLERRRVCENLPPDDELADHIGFOEAYRRFYGPV 49  
DB 1 YLYQWLGAAPVPPDPLERRRVCENLPPDDELADHIGFOEAYRRFYGPV 49

RESULT 2  
5434245-2  
;PATENT NO. 5434245  
; APPLICANT: KOYAMA, NOBUTO;KIMIZUKA, FUSAO;KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/993,980  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989  
; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989  
; SEQ ID NO:2:  
; LENGTH: 50  
5434245-2

Query Match 100.0%; Score 49; DB 6; Length 50;  
Best Local Similarity 100.0%; Pred. No. 5.9e-45;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILYQWLGAVPVPPDPLEPRREVCENPDCELADHIGFQEAIRRYGPV 499

DB 1 YLYQWLGAVPVPPDPLEPRREVCENPDCELADHIGFQEAIRRYGPV 499

RESULT 3  
 5434245-3  
 Patent No. 5434245  
 APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
 TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
 THE SAME  
 NUMBER OF SEQUENCES: 10  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/993,980  
 FILING DATE: 16-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 855,473; 444,786  
 FILING DATE: 23-MAR-1992; 01-DEC-1989  
 APPLICATION NUMBER: 444,786  
 FILING DATE: 01-DEC-1989  
 SEQ ID NO.: 3  
 LENGTH: 50  
 5434245-3

Query Match	100.0%;	Score 49;	DB 6;	Length 50;
Best Local Similarity	100.0%;	Pred. No. 5.9e-45;		
Matches 49;	Conservative 0;	Mismatches 0;	Indels	

QY 1 YLYQWLGAVPVYPDPLEPRREVCELNPDCELADHIGFQEA YRRRTGVP 49  
 |||||  
 2 YLYQWLGAVPVYPDPLEPRREVCELNPDCELADHIGFQEA YRRRTGVP 50  
 DB

RESULT 4  
 5434245-4  
 Patent No. 5434245  
 APPLICANT: KORYAMA, NOBUOTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
 TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
 THE SAME  
 NUMBER OF SEQUENCES: 10  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/993,980  
 FILING DATE: 16-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 855,473; 444,786  
 FILING DATE: 23-MAR-1992; 01-DEC-1989  
 APPLICATION NUMBER: 444,786  
 FILING DATE: 01-DEC-1989  
 SEQ ID NO.:4  
 LENGTH: 51  
 5434245-4

Query Match	100.0%;	Score 49;	DB 6;	Length 51
Best Local Similarity	100.0%;	Pred. No. 6e-45;		
Matches 49;	Conservative 0;	Mismatches 0;	Indels	

OY 1 YLYQWLTGAPVPPDPLEPRREVCELNPDCDELADHIGFQEAAYRRFYGPV 45  
|||||  
Db 2 YLYQWLTGAPVPPDPLEPRREVCELNPDCDELADHIGFQEAAYRRFYGPV 50

```

RESULT 5
5164483-1
; Patent No. 5164483
; APPLICANT: Takashi, Kurihara;Eiji, Taniyama;Sachio, Hirose
; TITLE OF INVENTION:  $\gamma$ -CARBOXYGLUTAMATE DERIVATIVE METHOD
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN
; USING THE SAME
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/575, 639
; FILING DATE: 31-AUG-1990
;

```

```

; SEQ ID NO:1:
;      LENGTH: 49
5164483-1

```

Query Match	51.0%;	Score 25;	DB 6;	Length 49;
Best Local Similarity	100.0%;	Pred. No. 1.1e-19;		
Matches 25; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      25 LNPDCDELADHIGFQEA YR RFGPV 49
         |||||
Db      25 LNPDCDELADHIGFQEA YR RFGPV 49
```

```

RESULT 6
5164483-2
; Patent No. 5164483
; APPLICANT: Takashi, Kurihara:Eiji, Tanigawa:Sachio, Hirose
; TITLE OF INVENTION: Y-CARBOXYGLUTAMATE DERIVATIVE METHOD
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN
; USING THE SAME
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/575,639
; FILING DATE: 31-AUG-1990
; SEQ ID NO.:2:
; LENGTH: 49
5164483-2

```

query Match	51.0%;	Score 25;	DB 6;	Length 49;
Best Local Similarity	100.0%;	Pred. No. 1.1e-19;		
Matches 25; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 25 LNPDCDELADHIGFQEA YRRFYGPV 45  
|||  
Db 25 LNPDCDELADHIGFQEA YRRFYGPV 45

RESULT 7  
 5164483-3  
 Patent No. 5164483  
 APPLICANT: Takasbi, Kurihara:Eiji, Taniyama:Sachio, Hirose  
 TITLE OF INVENTION:  $\gamma$ -CARBOXYGLUTAMATE DERIVATIVE METHOD  
 FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN  
 USING THE SAME  
 NUMBER OF SEQUENCES: 3  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/575,639  
 FILING DATE: 31-AUG-1990  
 SEQ ID NO.: 3  
 LENGTH: 110  
 5164483-3

Query Match	51.0%;	Score 25;	DB 6;	Length 98;
Best Local Similarity	100.0%;	Pred. No. 2.1e-19;		
Matches 25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 25 LNPDCDELADHIGFQEA YRRFFGPV 49  
 |||  
 Db 25 LNPDCDELADHIGFQEA YRRFFGPV 49

```

RESULT 8
5168041-1
; Patent No. 5168041
; APPLICANT: BERGMANN, ANDREAS E.
; TITLE OF INVENTION: METHOD FOR THE DETERMINATION OF
; OSTEOCALCIN IN HUMAN SERUM OR PLASMA
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/416,728
; FILING DATE: 03-OCT-1989
; SEQ ID NO.1:
; LENGTH: 48

```

5168041-1

Query Match 38.8%; Score 19; DB 6; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2,4e-13;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPPDLEPR 19  
|||||  
DB 1 YLYQWLGAPVPPDLEPR 19

RESULT 9  
US-08-796-598-13

Sequence 13, Application US/08796598

Patent No. 3627659

GENERAL INFORMATION:

APPLICANT: PATTERSON, DALE H.

APPLICANT: TARR, GEORGE E.

TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING

POLYMERS USING MASS SPECTROMETRY.

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Administrator - Testa, Hurwitz &amp;

STREET: High Street Tower, 125 High Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/796,598

FILING DATE: 07-FEB-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/446,055

FILING DATE: 19-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: FLYNN Esq., Kerry A.

REGISTRATION NUMBER: 33,693

REFERENCE/DOCKET NUMBER: SYP-115

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-796-598-13

Query Match 26.5%; Score 13; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAVPYPPDLEPR 19  
|||||  
DB 1 GAVPYPPDLEPR 13

RESULT 10  
US-08-447-175A-13

Sequence 13, Application US/08447175A

Patent No. 3669240

GENERAL INFORMATION:

APPLICANT: PATTERSON, DALE H.

TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING

TITLE OF INVENTION: POLYMERS WITH A STATISTICAL CERTAINTY USING MASS  
SPECTROMETRY.

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Administrator - Testa, Hurwitz &amp;

STREET: High Street Tower, 125 High Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/447,175A

FILING DATE: 19-MAY-1995

CLASSIFICATION: 422

ATTORNEY/AGENT INFORMATION:

NAME: RAUSCHENBACH, Kurt

REGISTRATION NUMBER: 40,137

REFERENCE/DOCKET NUMBER: SYP-114

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-447-175A-13

Query Match 26.5%; Score 13; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAVPYPPDLEPR 19  
|||||  
DB 1 GAVPYPPDLEPR 13

RESULT 11  
US-08-943-915-6

Sequence 6, Application US/08943915

Patent No. 5998170

GENERAL INFORMATION:

APPLICANT: Itoh, No. 5998170yuki

APPLICANT: Martin, Frank

TITLE OF INVENTION: A FIBROBLAST GROWTH FACTOR

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: 1840 DeHavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/943,915

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mazza, Richard J.

REGISTRATION NUMBER: 27,657  
REFERENCE/DOCKET NUMBER: A-469  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805.447.4112  
TELEFAX: 805.447.1090  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-943-915-6

Query Match 26.5%; Score 13; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVYPDPLEPR 19  
|||||  
DB 1 GAPVYPDPLEPR 13

RESULT 12  
US-08-881-037-112  
Sequence 112, Application US/08881037  
Patent No. 6080588  
GENERAL INFORMATION:  
APPLICANT: Glick, Gary D.  
TITLE OF INVENTION: DNA BINDING ANTIBODIES  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/881,037  
FILING DATE: 23-JUN-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/443,540  
FILING DATE: 18-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Konski, Antoinette F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: 203442110710  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 813-5600  
TELEFAX: (650) 494-0792  
TELEX:  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-881-037-112

Query Match 26.5%; Score 13; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVYPDPLEPR 19

DB 1 GAPVYPDPLEPR 13  
|||||

RESULT 13  
US-08-652-816A-43  
Sequence 43, Application US/08652816A  
Patent No. 5872215  
GENERAL INFORMATION:  
APPLICANT: Osbourn, JK  
APPLICANT: Allen, DJ  
APPLICANT: McCafferty, JG  
TITLE OF INVENTION: Specific binding members, materials and  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,816A  
FILING DATE: 23-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125579.4  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125579.8  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206318.9  
FILING DATE: 24-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206372.6  
FILING DATE: 23-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9525004.9  
FILING DATE: 07-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610824.6  
FILING DATE: 23-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/02240  
FILING DATE: 02-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/244,597  
FILING DATE: 01-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/33308  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-652-816A-43

Query Match 26.5%; Score 13; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 4.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVYPDPLEPR 19  
|||||

Db 23 GABVPPDPLEPR 35

## RESULT 14

US-09-142-974B-5  
; Sequence 5, Application US/09142974B  
; Patent No. 6451995  
; GENERAL INFORMATION:  
; APPLICANT: Cheung, Nai-Kong V.  
; APPLICANT: Laison, Steven M.  
; APPLICANT: Guo, Hong-Fen  
; APPLICANT: Rivlin, Ken  
; APPLICANT: Sadelain, Michel  
; TITLE OF INVENTION: Single Chain Fv Constructs of Anti-Ganglioside GD2  
; FILE REFERENCE: MSK.P-013-USNP  
; CURRENT APPLICATION NUMBER: US/09/142,974B  
; CURRENT FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: PCT/US97/04427  
; PRIOR FILING DATE: 1997-03-20  
; PRIOR APPLICATION NUMBER: 60/013,703  
; PRIOR FILING DATE: 1996-03-20  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: antibody tag  
US-09-142-974B-5

Query Match 14.3%; Score 7; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 PDPLEPR 19  
Db 7 PDPLEPR 13

RESULT 15  
US-08-466-120-2  
; Sequence 2, Application US/08466120  
; Patent No. 5869284  
; GENERAL INFORMATION:  
; APPLICANT: CAO, ET AL.  
; TITLE OF INVENTION: Retinoic Acid Receptor Epsilon  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,120  
; FILING DATE: June 6, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/07266  
; FILING DATE: 24 JUN 94  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FERRARO, GREGORY D.  
; REGISTRATION NUMBER: 36,134  
; REFERENCE/DOCKET NUMBER: 325800-354

## TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 433 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
US-08-466-120-2

Query Match 14.3%; Score 7; DB 2; Length 433;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WLCAPVP 11  
Db 4 WLCAPVP 10

RESULT 16  
PCT-US94-07266-2  
; Sequence 2, Application PC/TUS9407266  
; GENERAL INFORMATION:  
; APPLICANT: CAO, ET AL.  
; TITLE OF INVENTION: Retinoic Acid Receptor Epsilon  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/07266  
; FILING DATE: Concurrently  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; ATTORNEY/AGENT INFORMATION:  
; FILING DATE:  
; NAME: FERRARO, GREGORY D.  
; REGISTRATION NUMBER: 36,134  
; REFERENCE/DOCKET NUMBER: 325800-125  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 433 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
PCT-US94-07266-2

Query Match 14.3%; Score 7; DB 5; Length 433;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WLCAPVP 11  
Db 4 WLCAPVP 10

RESULT 17  
US-08-333-358-8  
; Sequence 8, Application US/08333358  
; Patent No. 5571696  
; GENERAL INFORMATION:  
; APPLICANT: EVANS Ph.D., RONALD M.  
; APPLICANT: MANGELSDORF Ph.D., DAVID J.  
; APPLICANT: ONG MS., ESTELITA S.  
; APPLICANT: ORO Ph.D., ANTHONY E.  
; APPLICANT: BORGMEYER Ph.D., UWE K.  
; APPLICANT: GIGUERE Ph.D., VINCENT MNM  
; TITLE OF INVENTION: NOVEL RECEPTORS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark  
; STREET: 444 So. Flower St., Suite 2000  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: US  
; ZIP: 90071-2921  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333.358  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/761.068  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Reiter Ph.D., Stephen E.  
; REGISTRATION NUMBER: 31192  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 440 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-333-358-8

Query Match 14.3%; Score 7; DB 1; Length 440;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WIGAPVP 11  
| | | | |  
DB 4 WIGAPVP 10

RESULT 18  
US-08-463-694-8  
; Sequence 8, Application US/08463694  
; Patent No. 5696233  
; GENERAL INFORMATION:  
; APPLICANT: EVANS Ph.D., RONALD M.  
; APPLICANT: MANGELSDORF Ph.D., DAVID J.  
; APPLICANT: ONG MS., ESTELITA S.  
; APPLICANT: ORO Ph.D., ANTHONY E.  
; APPLICANT: BORGMEYER Ph.D., UWE K.  
; APPLICANT: GIGUERE Ph.D., VINCENT MNM  
; APPLICANT: YAO Mr., TSO-PANG MNM  
; TITLE OF INVENTION: NOVEL RECEPTORS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark

STREET: 444 So. Flower St., Suite 2000  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: US  
; ZIP: 90071-2921  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463.694  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/761.068  
; FILING DATE: 17-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Reiter Ph.D., Stephen E.  
; REGISTRATION NUMBER: 31192  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 440 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-694-8

Query Match 14.3%; Score 7; DB 1; Length 440;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WIGAPVP 11  
| | | | |  
DB 4 WIGAPVP 10

RESULT 19  
US-08-694-501-8  
; Sequence 8, Application US/08694501  
; Patent No. 5710004  
; GENERAL INFORMATION:  
; APPLICANT: EVANS Ph.D., RONALD M.  
; APPLICANT: MANGELSDORF Ph.D., DAVID J.  
; APPLICANT: ONG MS., ESTELITA S.  
; APPLICANT: ORO Ph.D., ANTHONY E.  
; APPLICANT: BORGMEYER Ph.D., UWE K.  
; APPLICANT: GIGUERE Ph.D., VINCENT MNM  
; APPLICANT: YAO Mr., TSO-PANG MNM  
; TITLE OF INVENTION: NOVEL RECEPTORS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark  
; STREET: 444 So. Flower St., Suite 2000  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: US  
; ZIP: 90071-2921  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/694.501  
; FILING DATE: 07-AUG-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333.358

FILING DATE:  
APPLICATION NUMBER: US/07/761,068  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter Ph.D., Stephen E.  
REGISTRATION NUMBER: 31192  
REFERENCE/DOCKET NUMBER: P31 8936  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 440 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-694-501-8

Query Match 14.3%; Score 7; DB 1; Length 440;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WLGAPVP 11  
|||||  
Db 4 WLGAPVP 10

RESULT 20  
US-08-373-935-1  
Sequence 1, Application US/08373935  
Patent No. 5747661  
GENERAL INFORMATION:  
APPLICANT: Evans, Ronald M.  
APPLICANT: Mangelsdorf, David J.  
APPLICANT: Willy, Patricia J.  
TITLE OF INVENTION: IDENTIFICATION OF A DISTINCT  
TITLE OF INVENTION: RETINOID-RESPONSIVE PATHWAY AND USES THEREFOR  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
STREET: 444 South Flower Street, Suite 2000  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/373,935  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter, Stephen E.  
REGISTRATION NUMBER: 31,192  
REFERENCE/DOCKET NUMBER: P41 9894  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-546-4737  
TELEFAX: 619-546-4737  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-373-935-1

Query Match 14.3%; Score 7; DB 1; Length 447;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WLGAPVP 11  
|||||  
Db 4 WLGAPVP 10

RESULT 21  
US-08-968-747-7  
Sequence 7, Application US/08968747  
Patent No. 6060595  
GENERAL INFORMATION:  
APPLICANT: Scaglioni et al.  
TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968,747  
FILING DATE: 03-SEP-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 08472/705001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-968-747-7

Query Match 12.2%; Score 6; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12  
|||||  
Db 1 GAPVPY 6

RESULT 22  
US-08-493-071-28  
Sequence 28, Application US/08493071  
Patent No. 6127149  
GENERAL INFORMATION:  
APPLICANT: Hirai, Yohei  
APPLICANT: Koshida, Shogo  
APPLICANT: Oka, Yumiko  
TITLE OF INVENTION: MODIFIED EPIMORPHIN  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER  
STREET: 99 CANAL CENTER PLAZA, SUITE 300  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/493,071  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Price, Robert L.  
REGISTRATION NUMBER: 22,685  
REFERENCE/DOCKET NUMBER: 715-107  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-684-1111  
TELEFAX: 703-684-1124  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-493-071-28

Query Match 12.2%; Score 6; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPVP 12  
DB 1 GAVPVP 6

RESULT 23  
US-08-791-522-4  
Sequence 4, Application US/08791522  
Patent No. 5935817  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/791,522  
FILING DATE: Filed Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0193 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 118195  
US-08-791-522-4

Query Match 12.2%; Score 6; DB 2; Length 139;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11  
DB 31 LGAPVP 36

RESULT 24  
US-09-314-777-4  
Sequence 4, Application US/09314777  
Patent No. 6110686  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/314,777  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/791,522  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0193 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 118195  
US-09-314-777-4

Query Match 12.2%; Score 6; DB 3; Length 139;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11  
DB 31 LGAPVP 36

RESULT 25



US-08-457-176-2  
; Sequence 2, Application US/08457176  
; Patent No. 5591826  
; GENERAL INFORMATION:  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth W.  
; APPLICANT: de la Chappelle, Albert  
; TITLE OF INVENTION: Mutator Gene and Hereditary  
; TITLE OF INVENTION: No. 5591826-Polyposis Colorectal Cancer  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner, Birch, McKie, and Beckett  
; STREET: 1001 G Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/457,176  
; FILING DATE: 01-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/160295  
; FILING DATE: 02-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 01107.44900  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202.508.9100  
; TELEFAX: 202.508.9299  
; TELEX: 197430 BBMB UT  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 934 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: YES  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; US-08-457-176-2  
  
Query Match 12.2%; Score 6; DB 1; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 7 GAPVY 12  
|||||  
Db 614 GAPVY 619  
  
RESULT 26  
US-08-457-175-2  
; Sequence 2, Application US/08457175  
; Patent No. 5693470  
; GENERAL INFORMATION:  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth W.  
; APPLICANT: de la Chappelle, Albert  
; TITLE OF INVENTION: Mutator Gene and Hereditary  
; TITLE OF INVENTION: No. 5693470-Polyposis Colorectal Cancer  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner, Birch, McKie, and Beckett  
; STREET: 1001 G Street, N.W.

CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/457,175  
; FILING DATE: 01-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/160295  
; FILING DATE: 02-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 01107.44900  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202.508.9100  
; TELEFAX: 202.508.9299  
; TELEX: 197430 BBMB UT  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 934 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: YES  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; US-08-457-175-2  
  
Query Match 12.2%; Score 6; DB 1; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 7 GAPVY 12  
|||||  
Db 614 GAPVY 619  
  
RESULT 27  
US-08-709-784-1  
; Sequence 1, Application US/08709784  
; Patent No. 6048701  
; GENERAL INFORMATION:  
; APPLICANT: The Johns Hopkins University  
; TITLE OF INVENTION: Antibody Detection of Mismatch Repair  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Allegretti, Ltd.  
; STREET: 1001 G Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20001-4597  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/709,784  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/480,351

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.57434  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 934 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-709-784-1

Query Match 12.2%; Score 6; DB 3; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12  
Db 614 GAPVPY 619

RESULT 28  
US-09-651-656-3  
Sequence 3, Application US/09651656  
Patent No. 6340566  
GENERAL INFORMATION:  
APPLICANT: MCCUTHEN-MALONEY, SANDRA  
APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY  
TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE  
TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,  
TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES  
FILE REFERENCE: IL-10689  
CURRENT APPLICATION NUMBER: US/09/651,656  
CURRENT FILING DATE: 2000-08-29  
PRIOR APPLICATION NUMBER: 60/192,764  
PRIOR FILING DATE: 2000-03-28  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 934  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-651-656-3

Query Match 12.2%; Score 6; DB 4; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12  
Db 614 GAPVPY 619

RESULT 29  
US-09-650-855-3  
Sequence 3, Application US/09650855  
Patent No. 6365355  
GENERAL INFORMATION:  
APPLICANT: MCCUTHEN-MALONEY, SANDRA  
APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY  
TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA  
TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA  
FILE REFERENCE: IL-10284

CURRENT APPLICATION NUMBER: US/09/650,855  
CURRENT FILING DATE: 2000-08-29  
PRIOR APPLICATION NUMBER: 60/192,764  
PRIOR FILING DATE: 2000-03-28  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 934  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-650-855-3

Query Match 12.2%; Score 6; DB 4; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPY 12  
Db 614 GAPVPY 619

RESULT 30  
US-09-045-632-49  
Sequence 49, Application US/09045632  
Patent No. 6001575  
GENERAL INFORMATION:  
APPLICANT: HUGANIR, Richard L.  
APPLICANT: Dong, Hualing  
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
TITLE OF INVENTION: GRIP-RELATED MOLECULES  
NUMBER OF SEQUENCES: 105  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/045,632  
FILING DATE: 19-MAR-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/041,016  
FILING DATE: 19-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Corless, Peter F.  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 48147/1699-CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1050 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-045-632-49

Query Match 12.2%; Score 6; DB 3; Length 1050;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 PLEPRR 20  
Db 864 PLEPRR 869

RESULT 31  
US-09-045-632-50  
; Sequence 50, Application US/09045632  
; Patent No. 6001375  
; GENERAL INFORMATION:  
; APPLICANT: Haganir, Richard L.  
; APPLICANT: Dong, Hualing  
; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
; TITLE OF INVENTION: GRIP-RELATED MOLECULES  
; NUMBER OF SEQUENCES: 105  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/045,632  
; FILING DATE: 19-MAR-1998  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 60/041,016  
; FILING DATE: 19-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Corless, Peter F.  
; REGISTRATION NUMBER: 33,860  
; REFERENCE/DOCKET NUMBER: 48147/1699-CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1050 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-045-632-50

Query Match 12.2%; Score 6; DB 3; Length 1050;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PLEPRR 20  
|||||  
DB 864 PLEPRR 869

RESULT 32  
US-07-642-734C-5  
; Sequence 5, Application US/07642734C  
; Patent No. 5824513  
; GENERAL INFORMATION:  
; APPLICANT: Katz, L  
; APPLICANT: Donadio, S  
; APPLICANT: Mcalpine, J B  
; TITLE OF INVENTION: Recombinant DNA Method for Producing  
; TITLE OF INVENTION: Erythromycin Analogs  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Edward H. Gorman  
; STREET: Abbott Laboratories D377/Ab6D-2 One Abbott  
; STREET: Park Rd  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: US

ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/642,734C  
; FILING DATE: 17-JAN-91  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Danckerts, Andreas M  
; REGISTRATION NUMBER: 32652  
; REFERENCE/DOCKET NUMBER: 4952.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 708-937-9396  
; TELEFAX: 708-938-2623  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3170 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-642-734C-5

Query Match 12.2%; Score 6; DB 2; Length 3170;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LADHIG 37  
|||||  
DB 2884 LADHIG 2889

RESULT 33  
US-08-439-009A-5  
; Sequence 5, Application US/08439009A  
; Patent No. 6004787  
; GENERAL INFORMATION:  
; APPLICANT: Donadio, S  
; APPLICANT: Katz, L  
; APPLICANT: Mcalpine, J B  
; TITLE OF INVENTION: Method of Directing Biosynthesis of  
; TITLE OF INVENTION: Specific Polypeptides  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Steven F. Weinstein  
; STREET: Abbott Laboratories D377/Ab6D-2 One Abbott  
; STREET: Park Rd  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: US  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/439,009A  
; FILING DATE: 11-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Casuto, Dianne  
; REGISTRATION NUMBER: 40,943  
; REFERENCE/DOCKET NUMBER: 4952.US.D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847-938-3137  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3170 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-439-009A-5

Query Match 12.2%; Score 6; DB 3; Length 3170;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LADHIG 37  
|||||  
DB 2884 LADHIG 2889

RESULT 34  
US-08-991-789A-138  
Sequence 138, Application US/08991789A  
Patent No. 6225054  
GENERAL INFORMATION:  
APPLICANT: Frudakis, Tony N.  
Smith, John M.  
Reed, Steven G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 292  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed IP Law Group  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,789A  
FILING DATE: 11-Dec-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 138:  
US-08-991-789A-138

Query Match 10.2%; Score 5; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
|||||  
DB 3 QEAYR 7

RESULT 35  
US-09-062-451-138  
Sequence 138, Application US/09062451  
Patent No. 6344550  
GENERAL INFORMATION:  
APPLICANT: Frudakis, Tony N.  
Smith, John M.  
Reed, Steven G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 297  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/062,451  
FILING DATE: 04-APR-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.419C2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-09-062-451-138

Query Match 10.2%; Score 5; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43  
|||||  
DB 3 QEAYR 7

RESULT 36  
US-09-598-326-138  
Sequence 138, Application US/09598326  
Patent No. 6423496  
GENERAL INFORMATION:  
APPLICANT: Frudakis, Tony N.  
Smith, John M.  
Reed, Steven G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 247  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed Intellectual Property Law Group PLLC  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/598,326  
FILING DATE: 20-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419D1

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-09-598-326-138

Query Match          10.2% Score 5; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 QEAYR 43
   11111
Db 3 QEAYR 7

RESULT 37
US-08-247-809A-10
; Sequence 10, Application US/08247809A
; Patent No. 5569823
; GENERAL INFORMATION:
; APPLICANT: Peter H. Schreier; Klaus Stenzel; Gunter Adam;
; APPLICANT: Edgar Maiss
; TITLE OF INVENTION: DEOXYRIBONUCLEIC ACIDS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,809A
; FILING DATE: May 23, 1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 43 178 45.6 (Germany)
; FILING DATE: May 28, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: Bayer 9049-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-247-809A-10

Query Match          10.2% Score 5; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 PDPLE 17
   11111
Db 8 PDPLE 12
```

```

RESULT 38
US-08-711-728-10
; Sequence 10, Application US/08711728
; Patent No. 5973135
; GENERAL INFORMATION:
; APPLICANT: Peter H. Schreier; Klaus Stenzel; Gunter Adam;
; APPLICANT: Edgar Maiss
; TITLE OF INVENTION: DEOXYRIBONUCLEIC ACIDS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/711,728
; FILING DATE: 03-SEPT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/247,809
; FILING DATE: 23-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 43178456
; FILING DATE: 28-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: Bayer 9049.1-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-711-728-10

Query Match          10.2% Score 5; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 PDPLE 17
   11111
Db 8 PDPLE 12

RESULT 39
PCT-US93-05647-18
; Sequence 18, Application PCT/TUS9305647
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College and Eunice Kennedy Shriver
; TITLE OF INVENTION: HETEROGENEOUS PROTEOLIPID PEPTIDE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Millitia Drive
; CITY: Lexington
; STATE: MA
; ZIP: 02173
```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/05647  
FILING DATE: 19930610  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/896,704  
FILING DATE: JUNE 10, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: H092-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US93-05647-18

Query Match 10.2%; Score 5; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 WLGAP 9  
|||||  
DB 6 WLGAP 10

RESULT 40  
US-08-101-041A-4  
Sequence 4, Application US/08101041A  
Patent No. 5541101  
GENERAL INFORMATION:  
APPLICANT: SAJI, Fumitaka  
APPLICANT: AZUMA, Chihito  
APPLICANT: KIMURA, Tadashi  
TITLE OF INVENTION: ANTI-OXYTOCIN RECEPTOR ANTIBODIES AND  
METHODS FOR THEIR PRODUCTION  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/101,041A  
FILING DATE: 03-AUG-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-206854  
FILING DATE: 03-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 002258-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..18  
OTHER INFORMATION: /note= "Amino acids 102-119 of the  
oxytocin receptor polypeptide."  
US-08-101-041A-4

Query Match 10.2%; Score 5; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 RYGP 48  
|||||  
DB 3 RYGP 7

RESULT 41  
US-09-025-706-6  
Sequence 6, Application US/09025706  
Patent No. 5958874  
GENERAL INFORMATION:  
APPLICANT: Clark, Richard A  
APPLICANT: Greiling, Doris  
APPLICANT: Gailit, James  
TITLE OF INVENTION: RECOMBINANT FIBRONECTIN-BASED  
EXTRACELLULAR MATRIX FOR WOUND HEALING  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jaekle Fleischmann & Muegel, LLP  
STREET: 39 State Street  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14614-1310  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/025,706  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Braham, Susan J  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 87653,97R263  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-262-3640  
TELEFAX: 716-262-4133  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-025-706-6

Query Match 10.2%; Score 5; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 PRREV 22  
|||||

Db 7 PRREV 11

## RESULT 42

US-09-025-622-6

Sequence 6, Application US/09025622

Patent No. 6194378

## GENERAL INFORMATION:

APPLICANT: Clark, Richard A

TITLE OF INVENTION: FIBRONECTIN PEPTIDES-BASED EXTRACELLULAR

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jaekle Fleischmann &amp; Muegel, LLP

STREET: 39 State Street

CITY: Rochester

STATE: New York

COUNTRY: USA

ZIP: 14614-1310

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/025,622

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Braman, Susan J

REGISTRATION NUMBER: 34,103

REFERENCE/DOCKET NUMBER: 87653.97R270

TELEPHONE: 716-262-3640

TELEFAX: 716-262-4133

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-025-622-6

Query Match 10.2%; Score 5; DB 4; Length 18;

Best Local Similarity 100.0%; Pred. No. 58;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 PRREV 22

Db 7 PRREV 11

## RESULT 43

US-08-602-999A-125

Sequence 125, Application US/08602999A

Patent No. 6184205

## GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.

APPLICANT: KAY, Brian K.

APPLICANT: THORN, Judith M.

APPLICANT: QUILIAM, Lawrence A.

APPLICANT: DER, Channing J.

APPLICANT: FOWLKES, Dana M.

APPLICANT: RIDER, James E.

TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF

NUMBER OF SEQUENCES: 467

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie &amp; Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/602,999A

FILING DATE: 16-FEB-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-202

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

INFORMATION FOR SEQ ID NO: 125:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-602-999A-125

Query Match 10.2%; Score 5; DB 4; Length 20;

Best Local Similarity 100.0%; Pred. No. 63;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVP 11

Db 7 GAPVP 11

## RESULT 44

US-09-500-124-125

Sequence 125, Application US/09500124

Patent No. 6432920

## GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.

APPLICANT: KAY, Brian K.

APPLICANT: THORN, Judith M.

APPLICANT: QUILIAM, Lawrence A.

APPLICANT: DER, Channing J.

APPLICANT: FOWLKES, Dana M.

TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF

NUMBER OF SEQUENCES: 467

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie &amp; Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/500,124

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/602,999

FILING DATE: 16-FEB-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 125:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-500-124-125

Query Match 10.2%; Score 5; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAPVP 11  
DB 7 GAPVP 11

RESULT 45  
US-08-403-378B-12  
Sequence 12, Application US/08403378B  
Patent No. 5759991  
GENERAL INFORMATION:  
APPLICANT: TOHDOH, NAOKI  
APPLICANT: TOJO, SHIN-ICHIRO  
APPLICANT: KOTIMA, SHIN-ICHI  
APPLICANT: UKI, YASUYUKI  
APPLICANT: NISHIHARA, TOSHIO  
APPLICANT: FUKUSHIMA, NOBUYUKI  
APPLICANT: IRIE, TSUNEMASA  
APPLICANT: ONO, KEIICHI  
APPLICANT: AGUI, HIDEO  
APPLICANT: OIKA, KOSEI  
TITLE OF INVENTION: NEUROTROPIC PEPTIDE DERIVATIVES  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 PENNSYLVANIA AVENUE, NW  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20037-3202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,378B  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-124688  
FILING DATE: 27-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-080398  
FILING DATE: 30-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-280590  
FILING DATE: 27-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-333241  
FILING DATE: 21-DEC-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-243003  
FILING DATE: 12-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/758,043

FILING DATE: 12-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/873,764  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP93/01214  
FILING DATE: 27-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGART, WADDELL A  
REGISTRATION NUMBER: 24,861  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
TELEX: 6491103  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: rat  
STRAIN: Wistar  
TISSUE TYPE: hippocampal tissue of brain  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..22  
US-08-403-378B-12

Query Match 10.2%; Score 5; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPV 10  
DB 3 LGAPV 7

Search completed: December 4, 2002, 15:40:24  
Job time : 16 secs



```

AAR32937
ID AAR32937 standard; peptide; 49 AA.
XX
AC AAR32937;
XX
DT 05-JUL-1993 (first entry)
XX
DE 17, 21, 24, Gla human osteocalcin peptide.
XX
KM Gamma-carboxyglutamic acid; OS.
XX
OS Synthetic.
XX
FH Key
FT MISC-difference 17
FT /label= OTHER
FT /note= "OTHER= gamma-carboxyglutamic acid"
FT MISC-difference 21
FT /label= OTHER
FT /note= "OTHER= gamma-carboxyglutamic acid"
FT MISC-difference 24
FT /label= OTHER
FT /note= "OTHER= gamma-carboxyglutamic acid"
FT /note= "OTHER= gamma-carboxyglutamic acid"
XX
PN JP05032697-A.
XX
PD 09-FEB-1993.
XX
PF 31-JUL-1991; 91JP-0213251.
XX
PR 31-JUL-1991; 91JP-0213251.
XX
PA (TEIJU ) TEIJIN LTD.
XX
DR WPI; 1993-088665/11.
XX
DT Synthetic human osteocalcin for standard in determ. of natural
PT osteocalcin - prep. by introducing gamma-carboxy:glutamic acid
as fluorenyl protected gp.
XX
PS Claim 2; Page 2; 10pp; Japanese.
XX
CC The synthetic 17, 21, 24, gamma-carboxyglutamic acid form of human
CC osteocalcin (OS) was produced by introducing protected Gla. The
CC substance may be produced in high yield and is useful as standard
CC for the determination of human OS.
CC See also AAR32936.
XX
SQ Sequence 49 AA;
XX
Query Match 51.0%; Score 25; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 25 LNPDCDELADHIGFQEA YRRFGPV 49
DB 25 LNPDCDELADHIGFQEA YRRFGPV 49
XX
RESULT 14
AAM34264
ID AAM34264 standard; peptide; 49 AA.
XX
AC AAM34264;
XX
DT 23-APR-1998 (first entry)
XX
DE Glu17-osteocalcin peptide #2.
XX
KM Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;
XX osteoporosis; human.
XX
OS Homo sapiens.

```

```

XX
FH Key
FT Modified-site 21
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site 24
FT /note= "gamma-carboxyglutamic acid"
XX
PN WO9738309-A1.
XX
PD 16-OCT-1997.
XX
PF 10-APR-1997; 97WO-JP01246.
XX
PR 27-FEB-1997; 97JP-0043331.
XX 10-APR-1996; 96JP-0088608.
XX
PA (EISA ) EISAI CO LTD.
XX
PI Kimura T, Morimoto S, Sakakibara S;
XX
DR WPI; 1997-512875/47.
XX
DT Antibody specific for Glu17-osteocalcin, or its fragment - for
PT diagnosis of bone disorders such as osteoporosis
XX
PS Claim 3; Page 16-17; 28pp; Japanese.
XX
CC This sequence represents a human Glu17-osteocalcin peptide. This sequence
CC is recognised by the antibody of the invention. The antibody of the
CC invention is an anti-Glu17-osteocalcin antibody or its fragment, which
CC binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The
CC antibody can be used for the diagnosis of bone related disorders, such as
CC osteoporosis.
XX
SQ Sequence 49 AA;
XX
Query Match 51.0%; Score 25; DB 18; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 25 LNPDCDELADHIGFQEA YRRFGPV 49
DB 25 LNPDCDELADHIGFQEA YRRFGPV 49
XX
RESULT 15
AAM34266
ID AAM34266 standard; peptide; 49 AA.
XX
AC AAM34266;
XX
DT 23-APR-1998 (first entry)
XX
DE Gla21-osteocalcin peptide #2.
XX
KM Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;
XX osteoporosis; human.
XX
OS Homo sapiens.
XX
FH Key
FT Modified-site 17
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site 21
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site 24
FT /note= "gamma-carboxyglutamic acid"
XX
PN WO9738309-A1.
XX
PD 16-OCT-1997.
XX
PF 10-APR-1997; 97WO-JP01246.

```

XX 27-FEB-1997; 97JP-0043331.  
 PR 10-APR-1996; 96JP-0088608.  
 XX  
 XX (EISA) EISAI CO LTD.  
 PA  
 XX Kimura T, Morimoto S, Sakakibara S;  
 PI  
 XX WPI: 1997-512875/47.  
 DR  
 XX  
 PT Antibody specific for Glu17-osteocalcin, or its fragment - for  
 PT diagnosis of bone disorders such as osteoporosis  
 PS  
 PS Claim 4; Page 18; 28pp; Japanese.  
 XX  
 CC This sequence represents a human Glu21-osteocalcin peptide. This sequence  
 CC is recognised by the antibody of the invention. The antibody of the  
 CC invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
 CC binds to Glu17-osteocalcin, Glu21-osteocalcin or their fragments. The  
 CC antibody can be used for the diagnosis of bone related disorders, such as  
 CC osteoporosis.  
 CC  
 SQ Sequence 49 AA;

Query Match 51.0%; Score 25; DB 18; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-19;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LNPDCDELADHIGFOEAYRFGPV 49  
 ||||||||||||||||||||  
 Db 25 LNPDCDELADHIGFOEAYRFGPV 49

RESULT 16  
 AAR06630  
 ID AAR06630 standard; peptide; 21 AA.  
 XX  
 AC AAR06630;  
 XX  
 DT 09-JAN-1991 (first entry)  
 XX  
 DE Human osteocalcin fragment.  
 XX  
 KW Osteocalcin: musculoskeletal disorders; sandwich assay.  
 XX  
 PN WO9009587-A.  
 XX  
 PD 23-AUG-1990.  
 XX  
 PF 08-FEB-1990; 90WO-JP00155.  
 XX  
 PR 02-OCT-1989; 89JP-0255306.  
 PR 10-FEB-1989; 89JP-0030003.  
 XX  
 PA (TEIJ) TEIJIN KK.  
 PI  
 PI Hosoda K, Honda H, Kubota T, Masuno Y;  
 DR WPI: 1990-275233/36.  
 XX  
 PT Immunoassay of human osteocalcin in diagnostic specimens - using  
 PT sandwich assay with antibodies raised respectively to N-terminal  
 PT and C-terminal sites on osteocalcin  
 PS  
 PS Disclosure: fig 1; 80pp; Japanese.  
 XX  
 CC A monoclonal antibody (MAb), OST-N20, is raised to this N-terminal  
 CC peptide (bases 1-20) of human osteocalcin. Additional MAbs (OST-  
 CC C7 and OST-C15) are raised to two C-terminal osteocalcin peptide  
 CC fragments (comprising bases 43-49 and 35-49 respectively). A  
 CC sandwich immunoassay is then carried out and musculoskeletal dis-  
 CC orders can be diagnosed. See also AAR06631.  
 CC

SQ Sequence 21 AA;

Query Match 40.8%; Score 20; DB 11; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAAPVPPDPLEPRR 20  
 ||||||||||||||||  
 Db 1 YLYQWLGAAPVPPDPLEPRR 20

RESULT 17  
 AAW01681  
 ID AAW01681 standard; protein; 49 AA.  
 XX  
 AC AAW01681;  
 XX  
 DT 01-APR-1997 (first entry)  
 XX  
 DE Bone Gla protein.  
 DE  
 DE Bone Gla protein.  
 XX  
 KW BGP: bone gla protein; osteocalcin; Vitamin K-dependent protein;  
 KW bone matrix; therapy; diagnosis; assay; metabolic bone disease.  
 XX  
 OS Homo sapiens.  
 OS

Key Location/Qualifiers

FT Misc-difference 1 /note= "underlined in specification"  
 FT Misc-difference 3 /note= "underlined in specification"  
 FT Misc-difference 12 /note= "underlined in specification"  
 FT Misc-difference 12 /note= "underlined in specification"  
 FT Modified-site 21 /label= OTHER  
 FT Modified-site 21 /label= OTHER  
 FT Misc-difference 23..29 /note= "gamma-carboxyglutamic acid"  
 FT Modified-site 24 /note= "underlined in specification"  
 FT /label= OTHER  
 FT Misc-difference 42 /note= "gamma-carboxyglutamic acid"  
 FT Misc-difference 46 /note= "underlined in specification"  
 FT Misc-difference 46 /note= "underlined in specification"

DE4340597-A1.

01-JUN-1995.

29-NOV-1993; 93DE-4340597.

29-NOV-1993; 93DE-4340597.

(HENN-) HENNING BERLIN GMBH.  
 (BRAH-) BRAHMS DIAGNOSTICA GMBH.

Bergmann A, Weckermann R;

WPI: 1995-201516/27.

Determn. of osteocalcin in serum or plasma - with addn. of divalent  
 metal ions to inhibit decompn. of the protein, useful in therapy  
 and diagnosis of bone disease

Disclosure: Column 1; 9pp; German.

A method for determining concentration of osteocalcin in serum or plasma  
 is improved with addition of divalent metal ions to inhibit decomposition  
 of the protein. The method is useful in therapy and diagnosis of bone  
 disease. The present sequence is osteocalcin (a vitamin K-dependent  
 protein) that is a component of the bone matrix, or alternatively bone Gla  
 protein).

```

XX Sequence 49 AA:
SQ
Query Match 40.8%; Score 20; DB 16; Length 49;
Best Local Similarity 100.0%; Pred. No. 3.2e-14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPDPLEPR 20
   |||||
Db 1 YLYQWLGAAPVPPDPLEPR 20

RESULT 18
AAR06407
ID AAR06407 standard; peptide; 20 AA.
AC AAR06407;
XX
XX 20-MAY-1998 (first entry)
DT
XX Osteocalcin AA 1-19, peptide ABI.
DE
XX Osteocalcin; antibody; bone; absorption; metabolism.
KW
XX Synthetic.
OS
XX JP04225162-A.
PN
XX 14-AUG-1992.
PD
XX 27-DEC-1990; 90UP-0415242.
PF
XX 27-DEC-1990; 90UP-0415242.
PR
XX 27-DEC-1990; 90UP-0415242.
XX
XX (TEIJU ) TEIJU LTD.
PA
XX WPI; 1992-320370/39.
DR
XX
XX Antibody to human osteocalcin - obt'd. by immunising with
PT peptide of 1 to 19th radicals of N-terminal of human
PT osteocalcin
PS
XX Disclosure; fig 1; 8pp; Japanese.
XX
XX The sequences given in AAR06407-9 correspond to fragments of human
CC osteocalcin. These fragments were used in the production of an
CC antibody to human osteocalcin. The fragments were injected into an
CC animal and the resulting antibodies were obtained. The antibodies
CC can be used in the determination of the rate of bone absorption in
CC bone metabolism.
XX
SQ Sequence 20 AA:
Query Match 38.8%; Score 19; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e-13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPDPLEPR 19
   |||||
Db 2 YLYQWLGAAPVPPDPLEPR 20

RESULT 19
AAR06631
ID AAR06631 standard; protein; 15 AA.
AC AAR06631;
XX
XX 09-JAN-1991 (first entry)
DT
XX Human osteocalcin fragment.
DE
XX Osteocalcin; musculoskeletal disorders; sandwich assay.
KW

```

```

XX Key Location/Qualifiers
FH Peptide 9..15
FT
XX WO9009587-A.
PN
XX 23-AUG-1990.
PD
XX
XX 08-FEB-1990; 90WO-JP00155.
PF
XX 02-OCT-1989; 89JP-0255306.
PR
XX 10-FEB-1989; 89JP-0030003.
XX
XX (TEIJU ) TEIJU KK.
PA
XX Hosoda K, Honda H, Kubota T, Masuho Y;
PI
XX WPI; 1990-275233/36.
DR
XX
XX Immunassay of human osteocalcin in diagnostic specimens - using
PT sandwich assay with antibodies raised respectively to N-terminal
PT and C-terminal sites on osteocalcin
PT
XX Disclosure; fig 3; 80pp; Japanese.
PS
XX Monoclonal antibodies (MAbs), OST-C7 and OST-C15) are raised to
CC this osteocalcin peptide fragment (C-terminal bases 43-49) and
CC a shorter constituent of this (C-terminal bases 35-49) resp-
CC ectively. A further Mab (OST-N20) is raised to a fragment
CC comprising N-terminal bases 1-20. A sandwich immunoassay is
CC then carried out and musculoskeletal dis- orders can be diag-
CC nosed. See also AAR06630.
XX
XX Sequence 15 AA:
SQ
Query Match 28.6%; Score 14; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IGFOEAYRRFYGPV 49
   |||||
Db 2 IGFOEAYRRFYGPV 15

RESULT 20
AAO17841
ID AAO17841 standard; Protein; 68 AA.
XX
XX AAO17841;
AC
XX 20-AUG-2002 (first entry)
DT
XX
XX Peptide presentation method related vector encoded protein #2.
DE
XX Peptide presentation; host cell surface; antibody isolation;
KW epitope mapping; bacteria; EBT-II.
XX
XX Unidentified.
OS
XX WO200234906-A2.
PN
XX
XX 02-MAY-2002.
PD
XX
XX 25-OCT-2001; 2001WO-DE04009.
PF
XX 26-OCT-2000; 2000DE-1053224.
PR
XX
XX (UYGE-) UNIV GEORG AUGUST GOETTINGEN.
PA
XX Kolmar H, Christmann A, Wentzel A;
PI
XX WPI; 2002-435621/46.
XX
XX N-PSDB; AAL47111.
DR

```

XX Cell-surface presentation of peptides or proteins, useful e.g. for  
PT isolating monospecific antibodies, comprises expression of a fusion  
PT sequence with truncated Intimin, in bacteria -  
XX  
XX  
PS Example 1; Fig 2; 44pp; German.  
XX  
CC The present invention relates to a method for presenting peptides or  
CC proteins on the surface of host bacteria. The method can be used to  
CC isolate, from polyclonal mixtures, monospecific antibodies that bind  
CC selectively to a particular surface-bound polypeptide, or more generally  
CC any specific binding partners, and for epitope mapping. The present  
CC sequence is a vector encoded protein fragment described in the  
CC exemplification of the invention.  
XX  
XX Sequence 68 AA:  
QY  
Db 6 LGAPVPYPPDPLEPR 19  
14 LGAPVPYPPDPLEPR 27  
RESULT 21  
ID AAM04605 AAM04605 standard; peptide; 13 AA.  
XX  
XX AAM04605;  
XX  
XX 13-AUG-1997 (first entry)  
XX  
XX Osteocalcin 7-19 fragment for mass spectrometry analysis.  
DE  
XX  
XX Mass spectrometry; polymer analysis; biopolymer analysis.  
KM  
XX  
XX Synthetic.  
OS  
XX  
XX WO9636986-A1.  
PN  
XX  
XX 21-NOV-1996.  
PD  
XX  
XX 17-MAY-1996; 96WO-US07146.  
PE  
XX  
XX 19-MAY-1995; 95US-0447175.  
PR  
XX  
XX 19-MAY-1995; 95US-0446055.  
PA  
XX  
XX (PERS-) PERSEPTIVE BIOSYSTEMS INC.  
XX  
XX Patterson DH, Tarr GE;  
PI  
XX  
XX WPI: 1997-012308/01.  
DR  
XX  
XX Sequencing polymers, e.g. DNA, RNA, peptide nucleic acids, proteins,  
PT etc. - by obtaining mass to charge ratios of polymer fragments,  
PT pref. using mass spectrometer, and performing statistical analysis  
XX  
PS Example 2; Page 32; 86pp; English.  
XX  
XX A method of obtaining sequence information about a polymer (e.g. DNA,  
CC RNA, peptide nucleic acids, proteins, peptides and carbohydrates)  
CC comprising monomers of known mass has been claimed. The present  
CC sequence represents a fragment of osteocalcin (7-19), and was used as  
CC an example as a digestion before analysis by mass spectrometry,  
CC using this novel on-plate strategy. Total sequence information  
CC from a nine well digestion can be represented in a single digestion or  
CC it is often derived from two or more wells. The methods, apparatus and  
CC kit (claimed) can be used for the analysis of polymers, particularly  
CC biopolymers, e.g. DNA, RNA, peptide nucleic acids, proteins, peptides  
CC and carbohydrates. It provides a rapid, automated and cost effective  
CC sequencing of polymers, with a statistical certainty.

[illegible]



DT 22-JUN-2001 (first entry)  
 XX Bone Gla protein peptide SEQ ID NO:1096.  
 DE  
 XX  
 XX  
 KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidyl; maleimido group; amino;  
 XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200069900-A2.  
 XX  
 PD 23-NOV-2000.  
 XX  
 PF 17-MAY-2000; 2000WO-US13576.  
 XX  
 PR 17-MAY-1999; 99US-0134406.  
 PR 10-SEP-1999; 99US-0153406.  
 PR 15-OCT-1999; 99US-0159783.  
 XX  
 PA (CONF-) CONJUCHEM INC.  
 XX  
 PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
 DR WPI; 2001-112059/12.  
 XX  
 XX  
 PT Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity  
 PT  
 XX  
 PS Disclosure; Page 554; 733pp; English.  
 XX  
 XX The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.  
 CC  
 XX  
 SQ Sequence 13 AA;  
 XX  
 XX  
 Query Match 26.5%; Score 13; DB 22; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 37 GFGQAVRRFYGPV 49  
 ||||||||||||  
 DB 1 GFGQAVRRFYGPV 13  
 ||||||||||||  
 RESULT 26  
 AAB70172  
 ID AAB70172 standard; peptide; 13 AA.  
 XX  
 AC AAB70172;  
 XX  
 DT 26-APR-2001 (first entry)  
 XX  
 DE E tag.  
 XX  
 KW Antibody; sperm; S19; contraception.

XX  
 OS Synthetic.  
 OS  
 PN WO200107083-A1.  
 XX  
 PD 01-FEB-2001.  
 XX  
 PF 21-JUL-2000; 2000WO-US19843.  
 XX  
 PR 23-JUL-1999; 99US-0145512.  
 XX  
 PA (UYVI-) UNIV VIRGINIA PATENT FOUND.  
 XX  
 PI Herr JC, Norton EJ, Diekmann AB;  
 DR WPI; 2001-182730/18.  
 XX  
 XX  
 PT New recombinant antibody derivative of monoclonal antibody S19, useful  
 PT as a reagent for purifying or detecting human spermatozoa, as an active  
 PT ingredient of a spermstatic agent, or as a component of a spermicidal  
 PT contraceptive  
 XX  
 PS Example 1; Page 43; 48pp; English.  
 XX  
 XX The present invention relates to a recombinant antibody capable of  
 CC specifically binding to sperm agglutination antigen-1 (SAGA-1).  
 CC The recombinant antibody has two peptide fragments of the S19  
 CC antibody and the fragments are joined together by a linker.  
 CC The recombinant monoclonal antibodies are useful in a passive  
 CC immunity composition for contraception as they inhibit the ability of  
 CC sperm to fertilize an egg. The antibodies may be used e.g. as an  
 CC active ingredient of a spermstatic agent, or as a component of a  
 CC spermicidal contraceptive.  
 CC  
 XX  
 SQ Sequence 13 AA;  
 XX  
 XX  
 Query Match 26.5%; Score 13; DB 22; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 GAVPPYDPDLEPR 19  
 ||||||||||||  
 DB 1 GAVPPYDPDLEPR 13  
 ||||||||||||  
 RESULT 27  
 AAO15409  
 ID AAO15409 standard; peptide; 13 AA.  
 XX  
 AC AAO15409;  
 XX  
 DT 27-SEP-2002 (first entry)  
 XX  
 DE E tag - peptide tag sequence.  
 XX  
 KW Mammalian; nuclear receptor cofactor; CF7; CF8; metabolism regulation;  
 KW cell homeostasis; cell proliferation; differentiation;  
 KW pathological cellular aberration; cellular defence mechanism; E tag.  
 XX  
 OS Synthetic.  
 OS  
 PN WO200242322-A2.  
 XX  
 PD 30-MAY-2002.  
 XX  
 PF 21-NOV-2001; 2001WO-EP13548.  
 XX  
 PR 21-NOV-2000; 2000EP-0125524.  
 XX  
 PA (LION-) LION BIOSCIENCE AG.  
 XX  
 PI Jackson D, Casarl G, Suckow J;  
 XX

DR WPI: 2002-566559/60.  
XX  
XX Novel nuclear receptor cofactors; CF7 and CF8 for identifying  
PT modulators useful for inhibiting cellular function of cofactor and for  
PT treating metabolic disorders, immunological indications and hormonal  
PT dysfunctions  
XX  
XX Disclosure: Page 16; 68pp; English.  
PS  
XX The invention comprises the amino acid and coding sequences of two  
CC mammalian nuclear receptor cofactor proteins (CF7 and CF8). The CF7 and  
CC CF8 protein sequences of the invention are useful for screening agents  
CC that are capable of inhibiting the cellular function of cofactor CF7  
CC and/or CF8. The CF7 and CF8 DNA sequences are useful for modulating genes  
CC involved in cellular functions, such as: regulation of metabolism and  
CC cell homeostasis, cell proliferation and differentiation, pathological  
CC cellular aberrations, or cellular defence mechanisms. The present amino  
CC acid sequence represents a peptide tag that was used in the invention.  
XX  
SQ Sequence 13 AA;  
  
Query Match 26.5%; Score 13; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 7 GAPVPPDPLEPR 19  
DB 1 GAPVPPDPLEPR 13  
  
RESULT 28  
ABB83453  
ID ABB83453 standard; Peptide: 13 AA.  
XX  
XX ABB83453;  
AC  
XX  
XX 30-SEP-2002 (first entry)  
DT  
XX  
XX E tag.  
DE  
XX  
XX Nuclear receptor cofactor; CF11; metabolism regulation;  
KW cell homeostasis; cell proliferation; cell differentiation;  
KW central nervous system; metabolic disorder; immunological disease;  
KW hormonal dysfunction; neuronal disease.  
XX  
XX Synthetic.  
OS  
XX  
XX WO200253585-A2.  
PN  
XX  
XX 11-JUL-2002.  
PD  
XX  
XX 21-DEC-2001; 2001WO-EP15133.  
PF  
XX  
XX 30-DEC-2000; 2000EP-0128768.  
PR  
XX  
XX (LION-) LION BIOSCIENCE AG.  
PA  
XX  
XX Jackson D, Casari G, Suckow J;  
PI  
XX  
XX WPI: 2002-583601/62.  
DR  
XX  
XX Novel polypeptide useful for screening agents capable of inhibiting or  
PT activating the cellular function of cofactor of mammalian nuclear  
PT receptor CF11  
XX  
XX Disclosure: Page 15; 59pp; English.  
PS  
XX The present invention relates to CF11, a novel human nuclear receptor  
CC cofactor (ABB83453). CF11 modulates genes involved in various cellular  
CC functions such as regulation of metabolism and cell homeostasis, cell  
CC proliferation and differentiation, pathological cellular aberrations or  
CC cellular defence mechanisms. CF11 binds retinoic acid receptor (RAR;  
CC NR1A) and thyroid hormone receptor (TR; NR1B) but not retinoid X receptor

CC (TR; NR2B) or steroid hormone receptors. CF11 expression is restricted to  
CC the central nervous system and could be confined to neurons in the  
CC dentate gyrus of the hippocampus, the amygdala, thalamic and hypothalamic  
CC regions. Antagonists to CF11 are useful in the development of drugs  
CC against diseases such as metabolic disorders, immunological indications,  
CC hormonal dysfunctions and/or neurosystemic diseases or related to defects  
CC in neuronal diseases. The present sequence is a peptide tag, which may be  
CC used to generate recombinant CF11 proteins.  
XX  
SQ Sequence 13 AA;  
  
Query Match 26.5%; Score 13; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.6e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 7 GAPVPPDPLEPR 19  
DB 1 GAPVPPDPLEPR 13  
  
RESULT 29  
AAU98372  
ID AAU98372 standard; Peptide: 13 AA.  
XX  
XX AAU98372;  
AC  
XX  
XX 24-SEP-2002 (first entry)  
DT  
XX  
XX Synthetic E tag peptide sequence.  
DE  
XX  
XX Cofactor 9; CF9; nuclear receptor; metabolic disorder;  
KW drug design; immunological indication; hormonal dysfunction;  
KW neurosystemic disease; cofactor binding assay; E tag.  
XX  
XX Synthetic.  
OS  
XX  
XX WO200244365-A1.  
PN  
XX  
XX 06-JUN-2002.  
PD  
XX  
XX 28-NOV-2001; 2001WO-EP13891.  
PF  
XX  
XX 28-NOV-2000; 2000EP-0126022.  
PR  
XX  
XX (LION-) LION BIOSCIENCE AG.  
PA  
XX  
XX Jackson D, Casari G, Suckow J;  
PI  
XX  
XX WPI: 2002-527709/56.  
DR  
XX  
XX Novel mammalian nuclear receptor cofactor 9 polypeptide useful for  
PT identifying compounds for treating metabolic disorders, immunological  
PT indications, hormonal dysfunctions and/or neurosystemic diseases  
PT  
XX  
XX Disclosure: Page 16; 62pp; English.  
PS  
XX The present invention relates to a new mammalian nuclear receptor  
CC cofactor 9 (CF9) polypeptide. The invention is useful for screening for  
CC agents which are capable of inhibiting the cellular function of CF9.  
CC The invention is also useful for screening for nuclear receptors,  
CC for screening drugs for agonist and antagonist activity and for  
CC screening drugs useful in regulating physiological responses associated  
CC with CF9, and in structural drug designing. The molecules of the  
CC invention are useful for identifying compounds for treating  
CC metabolic disorders, immunological indications, hormonal dysfunctions  
CC and/or neurosystemic diseases. The invention can be useful as  
CC intermediates for making cellular preparations for cofactor binding  
CC assays, which are useful in drug screening. The present amino acid  
CC sequence represents the E tag peptide sequence, as described in the  
CC invention.  
XX  
SQ Sequence 13 AA;

Query Match	26.5%;	Score 13;	DB 23;	Length 13;
Best Local Similarity	100.0%;	Pred. No. 3.6e-07;		
Matches 13;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      7 GAPVPYPDPLEPR 19
        |||||
Db      1 GAPVPYPDPLEPR 13
```

RESULT 30
AAU78634
ID AAU78634 standard; Peptide; 13 AA.

AC AAU78634;

DT 18-JUN-2002 (first entry)

Synthetic E tag peptide sequence.

KW Pregnane X; receptor; PXR; cofactor; xenobiotic; E tag.

OS Synthetic.

PN WO200218420-A2.

PD 07-MAR-2002.

PF 17-AUG-2001; 2001WO-EP09488.

PR 28-AUG-2000; 2000EP-0118634.

PA (LION-) LION BIOSCIENCE AG.

PI Albers M, Ellwanger S, Koegl M, Loeser E;

DR WPI; 2002-292195/33

PT New nucleic acids and cofactors of the pregnane x nuclear receptor  
PT (PXR), which the nucleic acid encodes, useful for screening agonists or  
PT antagonists of PXR, and for determining a subject's response to  
PT xenobiotic substances or drugs -  
XX  
SS Disclosure, Page 17, 102pp; English.

CC The present invention relates to new nucleic acid molecule and its  
CC encoded polypeptide. The nucleic acid codes for a cofactor of the  
CC pregnane x nuclear receptor (PXR). The polypeptide encoded by the  
CC nucleic acid comprises 225 amino acids or 293 amino acids fully defined  
CC in the specification. The nucleic acid is useful for making vectors and  
CC transforming cells, both of which are ultimately useful for producing the  
CC cofactor proteins. The nucleic acids may also be used for determining a  
CC subject's response to xenobiotic substances or drugs. The proteins or  
CC complexes are useful for screening substances that bind the proteins or  
CC complexes, particularly agonists or antagonists of PXR. The present  
CC amino acid sequence represents the E tag peptide that was used in the  
CC invention to facilitate purification of recombinant proteins. The  
CC cofactor proteins of the invention are useful for screening for PXR.

**SQ Sequence 13 AA;**

Query Match	26.5%	Score 13;	DB 23;	Length 13;
Best Local Similarity	100.0%;	Pred. NO.	3.6e-07;	
Matches 13; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	7	GAPVPYPDPLEPR	19
Db	1	GAPVPYPDPLEPR	13

```
RESULT 31
AAE18828
ID AAE18828 standard; peptide; 13 AA.
XX
```

AC	AAE18828;
XX	
DT	17-MAY-2002 (first entry)

DE E tag used in the production of FGF-like polypeptides.

KM fibroblast growth factor; FGF-like protein wound healing; bullous  
KM edematomatous; erosive gastritis; inflammatory bowel disease; ulcer  
KM oesophagitis; Crohn's disease; hyaline membrane disease; emphysema;  
KM pulmonary fibrosis; hepatic cirrhosis; liver failure; angiodysplasia;  
KM multiple sclerosis; neurodegenerative disease; lung abnormality;  
KM viral hepatitis; respiratory distress syndrome; tumour; skin aging;  
KM gene therapy; vaccine; human; Etg.

OS Homo sapiens.

PN US2002001825-A1.

PD 03-JAN-2002.

PF 02-APR-2001; 2001US-0822485.

PR 31-MAR-2000; 2000US-0540118.

РА (ИТОН/) ИТОН N.

Itoh N;

DR WPI; 2002-187704/24.

PT Novel fibroblast growth factor-1-like polypeptide useful for treating,  
PT ameliorating and/or preventing dermal wounds, gastric ulcer, Crohn's  
PT disease and pulmonary inflammation -  
XX  
XX Example 3; Page 31; 63pp; English.

CC The invention relates to fibroblast growth factor (FGF)-like  
CC polypeptides and nucleic acid molecules encoding such polypeptides.  
CC Sequences of the invention are useful for treating, preventing or  
CC ameliorating a medical condition. They are useful for treating dermal  
CC wounds, eczematous, bullous, male pattern alopecia, gastric ulcer,  
CC duodenal ulcer, erosive gastritis, oesophagitis, oesophagogastric reflux  
CC disease, inflammatory bowel disease, Crohn's disease, radiation- or  
CC chemotherapy-induced gut toxicity, hyaline membrane disease, necrosis  
CC of the respiratory epithelium, emphysema, pulmonary inflammation,  
CC pulmonary fibrosis, hepatic cirrhosis, toxic insults to the liver,  
CC fulminant liver failure, viral hepatitis, mucositis, multiple sclerosis  
CC and other neurodegenerative diseases, infantile respiratory distress  
CC syndrome, bronchopulmonary dysplasia, acute respiratory distress  
CC syndrome or other lung abnormalities, tumours of the eye or the other  
CC tissues and organs. FGF-like polypeptides are useful stimulating  
CC angiogenesis, promoting wound healing, modulating differentiation of  
CC neuronal cells, adipocytes and skeletal muscle cells, preventing or  
CC ameliorate skin aging, preventing hair loss, stimulating the growth  
CC and differentiation of haematopoietic cells and bone marrow cells and  
CC maintaining organs before transplantation and for supporting cultures  
CC of primary cells and tissues. Sequences of the invention are also  
CC used in gene therapy and as vaccines. The present sequence is E tag  
CC used in the production of FGF-like polypeptides.

**SQ** Sequence 13 AA;

Query Match	26.5%	Score 13;	DB 23;	Length 13;
Best Local Similarity	100.0%	Pred. No. 3.6e-07;		
Matches 13; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	7	GAPVPYPDPLEPR	19
Db	1	GAPVPYPDPLEPR	13

RESULT 32  
AAE18389



```

ID  AAE18389 standard; peptide; 13 AA.
XX
AC  AAE18389;
XX
DT  07-MAY-2002 (first entry)
XX
DE  Human BGP peptide #1.
XX
KM  Human; cell proliferation; cell differentiation; parathyroid hormone;
KM  PTH; parathyroid related peptide; hyperproliferative skin disorder;
KM  psoriasis; ichthyosis; actinic keratosis; alopecia; skin cell growth;
KM  hair growth; wrinkle; wound healing; chemotherapy; skin cancer; PTHRP;
KM  bone Gla protein; BGP.
XX
OS  Homo sapiens.
XX
PN  WO200198348-A2.
XX
PD  27-DEC-2001.
XX
PF  20-JUN-2001; 2001WO-US19650.
XX
PR  22-JUN-2000; 2000US-213247P.
XX
PA  (HOLI/) HOLICK M F.
XX
PI  Holick MF;
XX
DR  WPI; 2002-171552/22.
XX
PT  Modulating proliferation or differentiation of mammalian skin or hair
PT  cell for treating hyperproliferative diseases, comprises topical
PT  administration of liposome encapsulated peptide identical to
PT  parathyroid hormone
XX
PS  Disclosure; Page 10; 58pp; English.
XX
CC  The invention relates to a method of modulating proliferation or
CC  differentiation of mammalian skin or hair cell. The method involves
CC  topical administration of a liposome encapsulated peptide, its salt
CC  or derivative, identical to parathyroid hormone (PTH) or parathyroid
CC  related peptide (PTHrP). Method of the invention is useful for
CC  inhibiting hyperproliferative skin disorders e.g. psoriasis,
CC  ichthyosis, actinic keratosis, skin cancer, inhibiting hair growth or
CC  preventing hair regrowth. It is also useful for stimulating skin cell
CC  growth, rejuvenating aged skin, preventing and treating skin wrinkles,
CC  enhancing wound healing, stimulating hair growth, maintaining hair
CC  growth, treating or preventing female or male pattern baldness,
CC  treating chemotherapy induced alopecia and for stimulating epidermal
CC  cell growth and hair follicle cell growth. The present sequence is
CC  a peptide of human bone Gla protein (BGP) used in the method of the
CC  invention.
XX
SQ  Sequence 13 AA;

Query Match      26.5%; Score 13; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  37 GFOEAYRRFYGPV 49
    |||||||||||
DB  1 GFOEAYRRFYGPV 13

RESULT 33
ID  AAM48056 standard; peptide; 13 AA.
XX
AC  AAM48056;
XX
DT  14-MAR-2002 (first entry)
XX
DE  E-tag sequence.

```

```

XX
KM  Phenotype; protein binding partner; ligand; E-tag.
XX
OS  Synthetic.
XX
PN  WO200186297-A2.
XX
PD  15-NOV-2001.
XX
PF  09-MAY-2001; 2001WO-US15092.
XX
PR  09-MAY-2000; 2000US-202912P.
XX
PA  (DGIB-) DGI BIOTECHNOLOGIES INC.
XX
PI  Blume AJ, Goldstein N, Pillutla R, Hsiao K, Prendergast J;
XX
DR  WPI; 2002-089808/12.
XX
PT  Identifying a naturally occurring binding partner or binding partner
PT  precursor for a target, comprises comparing identified amino acid
PT  sequence motifs to known sequences of a genome to identify a gene
PT  product of the genome having the motif
XX
PS  Example 1; Page 20; 47pp; English.
XX
CC  The invention relates to identifying a naturally occurring binding
CC  partner or binding partner precursor for a target, comprising comparing
CC  the identified amino acid sequence motifs to known amino acid sequences
CC  of a genome and identifying a gene product of the genome possessing the
CC  motif as the naturally occurring binding partner or partner precursor,
CC  for the target. The method is useful for determining the influence which
CC  specific genotypes have on phenotypes and for obtaining peptides which
CC  may be used to identify the natural protein partner of the target and
CC  enable synthesis of peptides which alter the phenotype of cells
CC  expressing the target. The method may also be used to identify peptide
CC  ligands capable of activating or inhibiting gene products through their
CC  ability to bind to such gene products as well as the activity and
CC  function of the gene products themselves. The present sequence is that of
CC  a tag sequence for peptide library sequences useful to the invention.
XX
SQ  Sequence 13 AA;

Query Match      26.5%; Score 13; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  7 GAVPYPDLEPR 19
    |||||||||||
DB  1 GAVPYPDLEPR 13

RESULT 34
ID  AAY70693 standard; peptide; 15 AA.
XX
AC  AAY70693;
XX
DT  18-JUL-2000 (first entry)
XX
DE  E. coli antibody derived peptide to isolate attractin specific scFv.
XX
KM  Attractin; immune response; macrophage; monocyte; T cell; cancer; scFv;
KM  immunostimulant; immunosuppressed patient; immunodeficiency syndrome;
KM  transplant; autoimmune disease; antibody; single chain variable fragment.
XX
OS  Escherichia coli.
XX
PN  WO200015651-A1.
XX
PD  23-MAR-2000.
XX
PF  14-SEP-1999; 99WO-US20948.

```

XX 14-SEP-1998; 98US-0100137.  
XX (DAND ) DANA FARBER CANCER INST INC.  
XX Duke-Cohan JS, Schlossman SF;  
XX WPI: 2000-271373/23.  
XX  
XX Isolated nucleic acids encoding human attractin polypeptides useful for  
XX enhancing immune responses -  
XX  
XX Example 7; Page 48; 120pp; English.  
XX  
XX The patent discloses four forms of human attractin polypeptides  
XX which enhance immune response by promoting macrophage and monocyte  
XX spreading in the presence of T cells. These include soluble attractin-1  
XX and -2 and membrane attractin-1 and -2. These various forms of attractin  
XX are encoded by alternatively spliced mRNA molecule transcribed  
XX from a single gene. The present sequence is a peptide downstream to  
XX kappa light chain of an E. coli antibody. Antibody directed against this  
XX region is used in an affinity column to isolate E. coli soluble single  
XX chain variable fragment (scFv) specific for attractin. Attractin can be  
XX used to enhance immune response  
XX in immunosuppressed patients such as those undergoing chemo- and  
XX radio-therapy treatment for cancer or those suffering from common  
XX variable immunodeficiency syndrome. The protein may also be used to  
XX screen modulators (agonists and antagonists) of immune responses  
XX which may also be used to regulate immune reactions. Attractin  
XX antibodies can be used to inhibit immune response in transplant  
XX recipients or patients afflicted with autoimmune disease.  
XX  
XX Sequence 15 AA:  
SQ  
Query Match 26.5%; Score 13; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.1e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAPVYPDPLEPR 19  
DB 1 GAPVYPDPLEPR 13  
RESULT 35  
AA05366  
ID AAY05366 standard; peptide: 18 AA.  
XX  
XX AAY05366;  
AC  
XX 30-JUN-1999 (first entry)  
DT  
XX  
XX HBV specific single stranded antibody fragment.  
DE  
XX  
XX Single stranded antibody; hepatitis B virus; HBV core protein;  
KW HBV infection; viral proliferation inhibitor; viral DNA synthesis;  
KW gene therapy.  
XX  
XX Mus sp.  
OS  
XX  
XX WO9911792-A1.  
PN  
XX 11-MAR-1999.  
PD  
XX 02-SEP-1998; 98WO-JP03921.  
PF  
XX 02-SEP-1997; 97JP-0237054.  
PR  
XX  
XX (HAYASHI) HAYASHI N.  
PA (SUMO ) SUMITOMO PHARM CO LTD.  
XX  
XX Hayashi N, Tohoh N, Yamamoto H, Yamamoto M;  
PI  
XX WPI: 1999-243623/20.  
DR

DR N-PSDB; AAX33934.  
XX  
XX Single-stranded antibody against hepatitis B virus core protein,  
PT applicable as (gene) therapeutic agents for treatment of hepatitis B  
PT viral infections  
XX  
XX Example 4; Page 68; 72pp; English.  
PS  
XX  
XX This sequence is the single-stranded antibody of the invention, that  
CC has the capability of binding to a hepatitis B virus (HBV) core protein.  
CC Therapeutic agents can be formulated with the antibody for treatment of  
CC HBV infections by stopping proliferation of the virus through inhibition  
CC of viral DNA synthesis, and the gene encoding the antibody can be applied  
CC as an agent to gene therapy.  
XX  
XX Sequence 18 AA:  
SQ  
Query Match 26.5%; Score 13; DB 20; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.8e-07;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAPVYPDPLEPR 19  
DB 4 GAPVYPDPLEPR 16  
RESULT 36  
AA08451  
ID AA08451 standard; Peptide: 38 AA.  
XX  
XX AA08451;  
AC  
XX 18-JUN-2002 (first entry)  
DT  
XX  
XX Insulin/insulin-like growth factor receptor-binding peptide #407.  
DE  
XX  
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
KW ophthalmological; insulin; receptor; gene therapy; diabetes;  
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
KW diabetic retinopathy; neurological diseases; stroke;  
KW diabetic neuropathy.  
XX  
XX  
XX Synthetic.  
OS  
XX  
XX WO200172771-A2.  
PN  
XX  
XX 04-OCT-2001.  
PD  
XX  
XX 29-MAR-2000; 2000WO-US08528.  
PF  
XX  
XX 29-MAR-2000; 2000WO-US08528.  
PR  
XX  
XX (DGTB-) DGT BIOTECHNOLOGIES LLC.  
PA (NOVO ) NOVO NORDISK AS.  
XX  
XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;  
PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;  
PI Hansen PH, Ravera M, Hsiao K;  
PI  
XX  
XX WPI: 2002-025774/03.  
DR  
XX  
XX Modulating insulin activity in mammalian cells, for treating e.g.  
PT diabetes and tumours, comprises using peptides that bind to insulin or  
PT insulin-like growth factor receptors -  
XX  
XX Example 28; Page 132; 390pp; English.  
PS  
XX  
XX The invention relates to a method of modulating insulin activity in  
CC mammalian cells by administering a peptide that binds the insulin  
CC receptor (IR). A composition containing a peptide, optionally expressed  
CC from gene therapy vectors, that binds to site 1 of IR and an insulin  
CC agonist are useful for treating diabetes. Also, peptides that are  
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are

CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
 CC receptor agonists are useful for treating neurological diseases,  
 CC including stroke and diabetic neuropathy. The peptides are also useful in  
 CC screening for compounds that bind to IR or IGF-1 receptor, potential  
 CC therapeutics and research reagents. AAU88034-AAU90957 represent IR  
 CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
 CC of the invention.

XX  
 XX Sequence 38 AA;  
 SQ

Query Match 26.5%; Score 13; DB 23; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 9, 1e-07;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPPVPPDPLEPR 19  
 Db 24 GAPPVPPDPLEPR 36  
 |||||

RESULT 37  
 AAU88442  
 ID AAU88442 standard; Peptide; 39 AA.  
 XX  
 XX AAU88442;  
 AC  
 XX  
 XX 18-JUN-2002 (first entry)  
 DT  
 XX  
 XX Insulin/insulin-like growth factor receptor-binding peptide #398.  
 DE  
 XX  
 XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
 KM ophthalmological; insulin; receptor; gene therapy; diabetes;  
 KM insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
 KM diabetic retinopathy; neurological diseases; stroke;  
 KM diabetic neuropathy.  
 KW  
 XX  
 XX Synthetic.  
 OS  
 XX  
 PN WO200172771-A2.  
 PD  
 XX 04-OCT-2001.  
 PD  
 XX 29-MAR-2000; 2000MO-US08528.  
 PF  
 XX 29-MAR-2000; 2000MO-US08528.  
 PR  
 XX (DGIB-) DGI BIOTECHNOLOGIES LLC.  
 PA (NOVO) NOVO NORDISK AS.  
 XX  
 XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;  
 PI Brissette R, Spezzler J, Cheng W, Ostergaard S, Mandelki WS;  
 PI Hansen PH, Ravera M, Hsiao K;  
 DR WPI; 2002-025774/03.  
 DR  
 XX  
 XX Modulating insulin activity in mammalian cells, for treating e.g.  
 PT diabetes and tumours, comprises using peptides that bind to insulin or  
 PT insulin-like growth factor receptors -  
 PT  
 XX  
 XX Example 28; Page 131; 390pp; English.  
 PS  
 XX  
 XX The invention relates to a method of modulating insulin activity in  
 CC mammalian cells by administering a peptide that binds the insulin  
 CC receptor (IR). A composition containing a peptide, optionally expressed  
 CC from gene therapy vectors, that binds to site 1 of IR and an insulin  
 CC agonist are useful for treating diabetes. Also, peptides that are  
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
 CC receptor agonists are useful for treating neurological diseases,  
 CC including stroke and diabetic neuropathy. The peptides are also useful in  
 CC screening for compounds that bind to IR or IGF-1 receptor, potential  
 CC therapeutics and research reagents. AAU88034-AAU90957 represent IR

CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
 CC of the invention.

XX  
 XX Sequence 39 AA;  
 SQ

Query Match 26.5%; Score 13; DB 23; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 9, 4e-07;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPPVPPDPLEPR 19  
 Db 25 GAPPVPPDPLEPR 37  
 |||||

RESULT 38  
 AAU19897  
 ID AAU19897 standard; Protein; 43 AA.  
 XX  
 XX AAU19897;  
 AC  
 XX  
 XX 07-DEC-1997 (first entry)  
 DT  
 XX  
 XX Vector pUC119EHIS-encoded polypeptide.  
 DE  
 XX  
 XX Carcinoembryonic antigen; CEA; human; antibody; scFv;  
 KM tumour marker; lung cancer; breast cancer; colon cancer;  
 KM adenocarcinoma; diagnosis; vector; plasmid pUC119EHIS.  
 KW  
 XX  
 XX Synthetic.  
 OS  
 XX  
 XX Key location/Qualifiers  
 FH Peptide 23..34  
 FT /label= E\_tag  
 FT Peptide 38..43  
 FT /label= His\_tag  
 XX  
 XX WO9720932-A1.  
 PN  
 XX  
 XX 12-JUN-1997.  
 PD  
 XX  
 XX 09-DEC-1996; 96MO-GB03043.  
 PF  
 XX  
 XX 11-OCT-1996; 96GB-0021295.  
 PR 07-DEC-1995; 95GB-0025004.  
 PR 23-MAY-1996; 96GB-0010824.  
 XX  
 XX (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
 PA  
 XX  
 XX Allen DJ, McCafferty JG, Osbourn JK;  
 PI  
 XX WPI; 1997-319779/29.  
 DR N-PSDB; AAT72137.  
 DR  
 XX  
 XX Specific binding members for human carcinoembryonic antigen - bind  
 PT to the A3-B3 extracellular domain of CEA and are substantially  
 PT non-cross-reactive with human liver cells; used for diagnosing  
 PT cancer  
 PT  
 XX  
 XX Example 7; Fig 11; 128pp; English.  
 PS  
 XX  
 XX This polypeptide sequence is encoded by the cloning site region  
 CC (see AAT72137) in vector pUC119EHIS. It includes an E tag for  
 CC detection with anti-E tag antibodies and a hexahistidine tag for  
 CC IMAC purification. Human carcinoembryonic antigen (CEA) epitope  
 CC sequences were cloned into the vector for preparation of purified  
 CC CEA domains. These were used to examine the domain recognition of  
 CC anti-CEA antibody CEA6 (see AAU19881) and T06D11.  
 CC  
 XX  
 XX Sequence 43 AA;  
 SQ

Query Match 26.5%; Score 13; DB 18; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1e-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAVPYPDPLEPR 19  
| | | | | | | | | |  
DB 23 GAVPYPDPLEPR 35

## RESULT 39

AAU88438  
ID AAU88438 standard; Peptide; 44 AA.

XX  
AC AAU88438;

XX  
DT 18-JUN-2002 (first entry)

XX  
DE Insulin/insulin-like growth factor receptor-binding peptide #394.

XX  
KW Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
KW ophthalmological; insulin; receptor; gene therapy; diabetes;  
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
KW diabetic retinopathy; neurological diseases; stroke;  
KW diabetic neuropathy.

XX  
OS Synthetic.

XX  
PN WO200172771-A2.

XX  
PD 04-OCT-2001.

XX  
PF 29-MAR-2000; 2000WO-US08528.

XX  
PR 29-MAR-2000; 2000WO-US08528.

XX  
PA (DGIB-) DGI BIOTECHNOLOGIES LLC.  
(NOVO) NOVO NORDISK AS.

XX  
PI Beasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;  
PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;

PI  
PI Hansen PH, Ravera M, Hsiao K;

XX  
DR WPI; 2002-025774/03.

XX  
PT Modulating insulin activity in mammalian cells, for treating e.g.  
PT diabetes and tumours, comprises using peptides that bind to insulin or  
PT insulin-like growth factor receptors -  
PS Example 28; Page 131; 390pp; English.

XX  
CC The invention relates to a method of modulating insulin activity in  
CC mammalian cells by administering a peptide that binds the insulin  
CC receptor (IR). A composition containing a peptide, optionally expressed  
CC from gene therapy vectors, that binds to site 1 of IR and an insulin  
CC agonist are useful for treating diabetes. Also, peptides that are  
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
CC receptor agonists are useful for treating neurological diseases,  
CC including stroke and diabetic neuropathy. The peptides are also useful in  
CC screening for compounds that bind to IR or IGF-1 receptor, potential  
CC therapeutics and research reagents. AAU88034-AAU90957 represent IR  
CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
CC of the invention.

XX  
SQ Sequence 44 AA;

Query Match 26.5%; Score 13; DB 23; Length 44;  
Best Local Similarity 100.0%; Pred. No. 1e-06; Indels 0; Gaps 0;  
Matches 13; Conservative 0; Mismatches 0;

OY 7 GAVPYPDPLEPR 19  
| | | | | | | | | |  
DB 30 GAVPYPDPLEPR 42

RESULT 40  
AAU88447  
ID AAU88447 standard; Peptide; 44 AA.

XX  
AC AAU88447;

XX  
DT 18-JUN-2002 (first entry)

XX  
DE Insulin/insulin-like growth factor receptor-binding peptide #403.

XX  
KW Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
KW ophthalmological; insulin; receptor; gene therapy; diabetes;  
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
KW diabetic retinopathy; neurological diseases; stroke;  
KW diabetic neuropathy.

XX  
OS Synthetic.

XX  
PN WO200172771-A2.

XX  
PD 04-OCT-2001.

XX  
PF 29-MAR-2000; 2000WO-US08528.

XX  
PR 29-MAR-2000; 2000WO-US08528.

XX  
PA (DGIB-) DGI BIOTECHNOLOGIES LLC.  
(NOVO) NOVO NORDISK AS.

XX  
PI Beasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;  
PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;

PI  
PI Hansen PH, Ravera M, Hsiao K;

XX  
DR WPI; 2002-025774/03.

XX  
PT Modulating insulin activity in mammalian cells, for treating e.g.  
PT diabetes and tumours, comprises using peptides that bind to insulin or  
PT insulin-like growth factor receptors -  
PS Example 28; Page 131; 390pp; English.

XX  
CC The invention relates to a method of modulating insulin activity in  
CC mammalian cells by administering a peptide that binds the insulin  
CC receptor (IR). A composition containing a peptide, optionally expressed  
CC from gene therapy vectors, that binds to site 1 of IR and an insulin  
CC agonist are useful for treating diabetes. Also, peptides that are  
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
CC receptor agonists are useful for treating neurological diseases,  
CC including stroke and diabetic neuropathy. The peptides are also useful in  
CC screening for compounds that bind to IR or IGF-1 receptor, potential  
CC therapeutics and research reagents. AAU88034-AAU90957 represent IR  
CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
CC of the invention.

XX  
SQ Sequence 44 AA;

Query Match 26.5%; Score 13; DB 23; Length 44;  
Best Local Similarity 100.0%; Pred. No. 1e-06; Indels 0; Gaps 0;  
Matches 13; Conservative 0; Mismatches 0;

OY 7 GAVPYPDPLEPR 19  
| | | | | | | | | |  
DB 30 GAVPYPDPLEPR 42

## RESULT 41

AAU88443  
ID AAU88443 standard; Peptide; 45 AA.

XX  
AC AAU88443;

XX

DT 18-JUN-2002 (first entry)  
XX Insulin/insulin-like growth factor receptor-binding peptide #399.  
XX  
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
KW ophthalmological; insulin: receptor; gene therapy; diabetes;  
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
KW diabetic retinopathy; neurological diseases; stroke;  
KW diabetic neuropathy.  
XX  
XX Synthetic.  
XX WO200172771-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 29-MAR-2000; 2000WO-US08528.  
XX  
XX 29-MAR-2000; 2000WO-US08528.  
XX  
XX (DGIB-) DGI BIOTECHNOLOGIES LLC.  
XX (NOVO) NOVO NORDISK AS.  
XX  
XX Beasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;  
PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;  
PI Hansen PH, Ravera M, Hsiao K;  
XX WPI: 2002-025774/03.  
XX  
XX Modulating insulin activity in mammalian cells, for treating e.g.  
PT diabetes and tumours, comprises using peptides that bind to insulin or  
PT insulin-like growth factor receptors -  
XX  
XX Example 28; Page 131; 390pp; English.  
XX  
XX The invention relates to a method of modulating insulin activity in  
CC mammalian cells by administering a peptide that binds the insulin  
CC receptor (R). A composition containing a peptide, optionally expressed  
CC from gene therapy vectors, that binds to Site 1 of IR and an insulin  
CC agonist are useful for treating diabetes. Also, peptides that are  
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
CC receptor agonists are useful for treating neurological diseases,  
CC including stroke and diabetic neuropathy. The peptides are also useful in  
CC screening for compounds that bind to IR or IGF-1 receptor, potential  
CC therapeutics and research reagents. AAU88034-AAU90957 represent IR  
CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
CC of the invention.  
XX  
XX Sequence 45 AA;  
SQ  
Query Match 26.5%; Score 13; DB 23; Length 45;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAVPYPPDLEPR 19  
DB 31 GAVPYPPDLEPR 43  
IIIIIIIIIIIIIIIIIIII  
RESULT 42  
AAU88440  
ID AAU88440 standard; Peptide: 46 AA.  
XX  
XX AAU88440;  
AC  
XX  
XX 18-JUN-2002 (first entry)  
DT  
XX Insulin/insulin-like growth factor receptor-binding peptide #396.  
DE  
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
KW ophthalmological; insulin: receptor; gene therapy; diabetes;  
KW ophthalmological; insulin: receptor; gene therapy; diabetes;  
XX

KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
KW diabetic retinopathy; neurological diseases; stroke;  
KW diabetic neuropathy.  
XX  
XX Synthetic.  
XX WO200172771-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 29-MAR-2000; 2000WO-US08528.  
XX  
XX 29-MAR-2000; 2000WO-US08528.  
XX  
XX (DGIB-) DGI BIOTECHNOLOGIES LLC.  
XX (NOVO) NOVO NORDISK AS.  
XX  
XX Beasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;  
PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;  
PI Hansen PH, Ravera M, Hsiao K;  
XX WPI: 2002-025774/03.  
XX  
XX Modulating insulin activity in mammalian cells, for treating e.g.  
PT diabetes and tumours, comprises using peptides that bind to insulin or  
PT insulin-like growth factor receptors -  
XX  
XX Example 28; Page 131; 390pp; English.  
XX  
XX The invention relates to a method of modulating insulin activity in  
CC mammalian cells by administering a peptide that binds the insulin  
CC receptor (R). A composition containing a peptide, optionally expressed  
CC from gene therapy vectors, that binds to Site 1 of IR and an insulin  
CC agonist are useful for treating diabetes. Also, peptides that are  
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
CC receptor agonists are useful for treating neurological diseases,  
CC including stroke and diabetic neuropathy. The peptides are also useful in  
CC screening for compounds that bind to IR or IGF-1 receptor, potential  
CC therapeutics and research reagents. AAU88034-AAU90957 represent IR  
CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
CC of the invention.  
XX  
XX Sequence 46 AA;  
SQ  
Query Match 26.5%; Score 13; DB 23; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAVPYPPDLEPR 19  
DB 32 GAVPYPPDLEPR 44  
IIIIIIIIIIIIIIIIIIIIII  
RESULT 43  
AAU88452  
ID AAU88452 standard; Peptide: 55 AA.  
XX  
XX AAU88452;  
AC  
XX  
XX 18-JUN-2002 (first entry)  
DT  
XX Insulin/insulin-like growth factor receptor-binding peptide #408.  
DE  
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
KW ophthalmological; insulin: receptor; gene therapy; diabetes;  
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
KW diabetic retinopathy; neurological diseases; stroke;  
KW diabetic neuropathy.  
XX  
XX Synthetic.  
OS

[illegible]

PR	29-MAR-2000; 2000OWO-US08528.
XX	(DGI B-) DGI BIOTECHNOLOGIES LLC.
PA	(NOVO ) NOVO NORDISK AS.
XX	Beeasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;
PI	Brissette J, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;
PI	Hansen PH, Ravera M, Hsiao K;
XX	WPI: 2002-025774/03.
DR	
XX	
PT	Modulating insulin activity in mammalian cells, for treating e.g.
PT	diabetes and tumours, comprises using peptides that bind to insulin or
PT	insulin-like growth factor receptors
XX	
PS	Claim 219, Page 182; 390pp: English.
XX	
CC	The invention relates to a method of modulating insulin activity in
CC	mammalian cells by administering a peptide that binds the insulin
CC	receptor (IR). A composition containing a peptide, optionally expressed
CC	from gene therapy vectors, that binds to site 1 of IR and an insulin
CC	agonist are useful for treating diabetes. Also, peptides that are
CC	antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
CC	useful for treating insulin-like growth factor (IGF)-sensitive tumours
CC	(e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
CC	receptor agonists are useful for treating neurological diseases,
CC	including stroke and diabetic neuropathy. The peptides are also useful in
CC	screening for compounds that bind to IR or IGF-1-receptor, potential
CC	therapeutics and research reagents. AAU88034-AAU90957 represent IR
CC	and/or IGF-1 receptor-binding peptides and related amino acid sequences
CC	of the invention.
XX	
SQ	Sequence 55 AA:
	Query Match 26.5%; Score 13; DB 23; Length 55;
	Best Local Similarity 100.0%; Pred. No. 1.3e-06;
	Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	7 GAPVPPDPLEPR 19 
DB	41 GAPVPPDPLEPR 53
RESULT 45	
AAU88454	
ID	AAU88454 standard; Peptide; 58 AA.
XX	
AC	AAU88454;
XX	
DT	18-JUN-2002 (first entry)
XX	
DE	Insulin/insulin-like growth factor receptor-binding peptide #410.
XX	
KW	Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
KW	ophthalmological; insulin; receptor; gene therapy; diabetes;
KW	insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
KW	diabetic retinopathy; neurological diseases; stroke;
KW	diabetic neuropathy.
XX	
OS	Synthetic.
XX	
PN	WO200172771-A2.
PD	04-OCT-2001.
XX	
PF	29-MAR-2000; 2000OWO-US08528.
XX	
PR	29-MAR-2000; 2000OWO-US08528.
XX	
PA	(DGI B-) DGI BIOTECHNOLOGIES LLC.
PA	(NOVO ) NOVO NORDISK AS.
XX	
PI	Beeasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;

PI Brissette R, Speizer J, Cheng W, Ostergaard S, Mandeckl WS;  
PI Hansen PH, Ravera M, Hsiao K;  
XX  
DR WPI; 2002-025774/03.

XX  
PT Modulating insulin activity in mammalian cells, for treating e.g.  
PT diabetes and tumours, comprises using peptides that bind to insulin or  
PT insulin-like growth factor receptors  
XX

PS Example 28; Page 132; 390pp; English.

XX  
CC The invention relates to a method of modulating insulin activity in  
CC mammalian cells by administering a peptide that binds the insulin  
CC receptor (IR). A composition containing a peptide, optionally expressed  
CC from gene therapy vectors, that binds to Site 1 of IR and an insulin  
CC agonist are useful for treating diabetes. Also, peptides that are  
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
CC receptor agonists are useful for treating neurological diseases,  
CC including stroke and diabetic neuropathy. The peptides are also useful in  
CC screening for compounds that bind to IR or IGF-1 receptor, potential  
CC therapeutics and research reagents. AA088034-AA090957 represent IR  
CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
CC of the invention.  
XX

SQ Sequence 58 AA;

Query Match 26.5%; Score 13; DB 23; Length 58;  
Best Local Similarity 100.0%; Pred. No. 1.3e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAVPYPPDPLEPR 19  
Db 44 GAVPYPPDPLEPR 56

Search completed: December 4, 2002, 15:38:42  
Job time : 37 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:29:49 : Search time 28 seconds  
(without alignments)  
360.582 Million cell updates/sec

Title: US-09-462-931-2  
Perfect score: 284  
Sequence: 1 XLYQWIGAPVPPDLEPRR.....DELADHIGFQEAHYRFGPV 49

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp rodent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*
- 15: sp.virus:\*
- 16: sp.bacteriap:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	65.1	34	6	Q9TT28
2	162	57.0	95	11	Q99K39
3	89	31.3	97	13	Q90VM2
4	76	26.8	147	11	Q925S3
5	76	26.8	170	11	Q925S2
6	76	26.8	298	11	Q9QYF0
7	74	26.1	104	13	Q9W668
8	71	25.0	441	10	Q8RXX0
9	63	22.2	1865	10	Q9C8G1
10	63	22.2	2254	10	Q38970
11	63	22.2	2254	10	Q38971
12	63	22.2	2257	10	Q9SKY1
13	62.5	22.0	10917	2	Q9JNW6
14	62	21.8	574	5	Q09967
15	61.5	21.7	609	16	Q92KR7
16	61.5	21.7	691	2	Q93G15

17	61	21.5	1034	4	Q96MS0	Q96MS0 homo sapien
18	60.5	21.3	677	16	Q8X7E7	Q8X7E7 escherichia
19	60	21.1	377	11	Q9EQ73	Q9EQ73 mus musculus
20	60	21.1	563	13	Q980H6	Q980H6 cyprinus ca
21	60	21.1	870	3	Q14424	Q14424 coccidioides
22	60	21.1	1344	11	Q92214	Q92214 mus musculus
23	60	21.1	2359	10	Q9C8G0	Q9C8G0 arabidopsis
24	60	21.1	2375	10	Q9FR96	Q9FR96 arabidopsis
25	59.5	21.0	369	4	Q43413	Q43413 homo sapien
26	59.5	21.0	458	10	Q9SXD5	Q9SXD5 arabidopsis
27	59.5	21.0	477	17	Q8ZSN7	Q8ZSN7 pyrobaculum
28	59.5	21.0	1188	4	Q9Q026	Q9Q026 homo sapien
29	59	20.8	706	5	Q18313	Q18313 caenorhabd
30	58.5	20.6	155	4	Q9H433	Q9H433 homo sapien
31	58.5	20.6	362	4	Q9HBW7	Q9HBW7 homo sapien
32	58.5	20.6	362	4	Q96P71	Q96P71 homo sapien
33	58.5	20.6	432	5	Q21857	Q21857 caenorhabd
34	58.5	20.6	490	2	Q51452	Q51452 pseudomonas
35	58.5	20.6	490	16	Q51460	Q51460 pseudomonas
36	58.5	20.6	699	16	Q8ZCD5	Q8ZCD5 yersinia pe
37	58	20.4	196	2	Q8VYH2	Q8VYH2 mesorhizobi
38	58	20.4	796	16	Q8XPG5	Q8XPG5 raietonia s
39	58	20.4	928	17	Q8ZVB8	Q8ZVB8 pyrobaculum
40	58	20.4	2267	10	Q8S6N5	Q8S6N5 oryza sativ
41	57.5	20.2	409	2	Q54763	Q54763 synecococc
42	57.5	20.2	412	5	Q9UAC5	Q9UAC5 dirosophila
43	57.5	20.2	486	5	Q9XYR0	Q9XYR0 dirosophila
44	57	20.1	41	16	Q9AD27	Q9AD27 streptomyces
45	57	20.1	187	2	Q8RKJ9	Q8RKJ9 bradyrhizob

## ALIGNMENTS

## RESULT 1

ID Q9TT28 PRELIMINARY: PRT: 34 AA.

AC Q9TT28: 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE Bone gamma-carboxylglutamate protein osteocalcin (Fragment).  
 GN GLA  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRED BEAGLE;  
 RA Wezeman F.H., Moskal S.F. II.;  
 RT "Canine Osteoblast Gene Expression."  
 RT Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF205942; AAF18456.1; -  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; VILK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA; 1.  
 FT NON\_TER 1  
 FT NON\_TER 34  
 SQ SEQUENCE 34 AA: 3807 MW: 3F323F8D6FDC65D8 CRC64:

Query Match 65.1%; Score 185; DB 6; Length 34;  
 Best Local Similarity 94.1%; Pred. No. 1.3e-16;  
 Matches 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 AVPPYDPLEPRREVCGLNPDDELADHIGFQEA 41  
 |||  
 DB 1 AVPPYDPLEPRREVCGLNPDDELADHIGFQEA 34

## RESULT 2

```
Q99K39
ID Q99K39 PRELIMINARY; PRT; 95 AA.
AC Q99K39;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to bone gamma-carboxylated protein, related sequence
DE 1.
GN A1461847.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC005483; AAB05483.1; -.
DR MGD; MGI:2139729; A1461847.
DR InterPro; IPR002384; GLA_bone.
DR InterPro; IPR000294; VltK_dep_GLA.
DR Pfam; PF00594; gla; 1.
DR PRINTS; PRO0002; GLABONE.
DR SMART; SM00069; GLA; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
SQ SEQUENCE 95 AA; 10445 MW; 25C4B3AAB851909E CRC64;

Query Match 57.0%; Score 162; DB 11; Length 95;
Best Local Similarity 63.0%; Pred. No. 3,4e-13;
Matches 29; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

OY 2 LYWGAGVPPDLEPRREVCELPDCELDADHIGFOEAYRRFYG 47
DB 47 LRRYLGASVSPDPLEPRCELDPADELSNQYGLKAYRRFYG 92

RESULT 3
OY Q90VW2 PRELIMINARY; PRT; 97 AA.
ID Q90VW2;
AC Q90VW2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Osteocalcin.
OS Sparus aurata (gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_taxid=8175;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21297182; PubMed=11404005;
RA Pinto J.P., Ohresser M.C.P., Canceila M.L.;
RA "Cloning of the bone gla protein gene from the teleost fish Sparus
RA aurata. Evidence for overall conservation in gene organization and
RA bone-specific expression from fish to man.";
RL Gene 270:77-91(2001).
DR EMBL; AF048703; AAK6568.1; -.
DR EMBL; AF289506; AAK62679.1; -.
DR InterPro; IPR000294; VltK_dep_GLA.
DR Pfam; PF00594; gla; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; UNKNOWN.1.
SQ SEQUENCE 97 AA; 10434 MW; B95608824FDFECB CRC64;

Query Match 31.3%; Score 89; DB 13; Length 97;
Best Local Similarity 50.0%; Pred. No. 0.00073;
Matches 17; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

OY 16 LEPRREVEINPDCELDADHIGFOEAYRRFYGPV 49
DB 62 LESIREVCELINACEHMDTEGIIAAYAYYGPV 95
```

```
RESULT 4
OY Q925S3 PRELIMINARY; PRT; 147 AA.
ID Q925S3;
AC Q925S3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE MRP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/C;
RC PubMed=11819679;
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,
RA Su C.;
RL "Mechanism of exogenous nucleic acids and their precursors improving
RL the repair of intestinal epithelium after irradiation in mice.";
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/C;
RC Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,
RA Su C.;
RL "Cloning of mouse genes related to repairing of intestinal epithelium
RL of the irradiated mice by treatment with the intestinal RNA of mice of
RL the same strain.";
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).
DR EMBL; AF240166; AAK43731.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
SQ SEQUENCE 147 AA; 16274 MW; 800594A12B97191F CRC64;

Query Match 26.8%; Score 76; DB 11; Length 147;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAVPYPPDLEPR 19
DB 133 GAVPYPPDLEPR 145
```

```
RESULT 5
OY Q925S2 PRELIMINARY; PRT; 170 AA.
ID Q925S2;
AC Q925S2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE MRP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/C;
RC PubMed=11819679;
RX Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,
RA Su C.;
RL "Mechanism of exogenous nucleic acids and their precursors improving
RL the repair of intestinal epithelium after irradiation in mice.";
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/C;
RC Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,
RA Su C.;
RL "Cloning of mouse genes related to repairing of intestinal epithelium
RL of the irradiated mice by treatment with the intestinal RNA of mice of
RL the same strain.";
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).
```

DR EMBL: AF240167; AAK43732.1; .  
 DR InterPro: IPR003598; Ig\_C2.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00408; IgC2; 1.  
 DR Immunoglobulin domain.  
 SO SEQUENCE 170 AA; 17978 MW; 5042823CC6C10F38 CRC64;

Query Match 26.8%; Score 76; DB 11; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 0.062;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 GAVPYPPDPLEPR 19  
 Db 156 GAVPYPPDPLEPR 168

## RESULT 6

O90YFO PRELIMINARY; PRT; 298 AA.  
 AC O90YFO;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE CN 8 scfv.  
 CN CN 8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-SPLEEN;  
 RA MEDLINE-20183931; PubMed-10706631;  
 RA Shinohara N.; Demura T.; Fukuda H.;  
 RT "Isolation of a vascular cell wall-specific monoclonal antibody  
 RT recognizing a cell polarity by using a phage display subtraction  
 RT method".  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).  
 DR EMBL: AB036341; BAA8633.1; .  
 DR HSSP: P01607; IREI.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 2.  
 DR SMART: SM00406; IgV; 2.  
 SO SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 26.8%; Score 76; DB 11; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 GAVPYPPDPLEPR 19  
 Db 284 GAVPYPPDPLEPR 296

## RESULT 7

O9W668 PRELIMINARY; PRT; 104 AA.  
 AC O9W668;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Matrix Gla protein.  
 GN MCP.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodidae; Xenopus.  
 OX NCBI\_TaxID=8335;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21429828; PubMed-11550673;

RA Cancelli M.L.; Ohresser M.C.; Reia J.P.; Viegas C.S.; Williamson M.K.,  
 RA Price P.A.;  
 RT "Matrix Gla protein in Xenopus laevis: molecular cloning, tissue  
 RT distribution, and evolutionary considerations."  
 RL J. Bone Miner. Res. 16:1611-1621(2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA Cancelli M.L.; Price P.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF055588; AAD28354.2; .  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; VltK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 2.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA; 2.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; UNKNOWN.2.  
 SO SEQUENCE 104 AA; 12326 MW; 46ABB9E3BC23D04A CRC64;

Query Match 26.1%; Score 74; DB 13; Length 104;  
 Best Local Similarity 42.4%; Pred. No. 0.064;  
 Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Oy 15 PLEPRREVCENLPPCDELADHIGFQEAAYRRFG 47  
 Db 65 PRERQRECEIDPCERIALRYGFTAAVKRYFG 97

## RESULT 8

O8RYX0 PRELIMINARY; PRT; 441 AA.  
 AC O8RYX0;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE OSJNBA0066C06.7 protein.  
 GN OSJNBA0066C06.7.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. NIPPONBARE;  
 RA Sasaki T.; Matsumoto T.; Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
 RT clone:OSJNBA0066C06."  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003933; BAB90635.1; .  
 SO SEQUENCE 441 AA; 47810 MW; 5B66364PF689AE13 CRC64;

Query Match 25.0%; Score 71; DB 10; Length 441;  
 Best Local Similarity 34.7%; Pred. No. 0.77;  
 Matches 17; Conservative 4; Mismatches 20; Indels 8; Gaps 1;

Oy 9 PVYPPDLEPRREVCENLPPCDELA-----DHIGFQEAAYRRFGPV 49  
 Db 251 PLPLPPPPPPPPRYTRSRDSSAATAGKTRLDHIGFEDLRRYFYMPI 299

## RESULT 9

O9C8G1 PRELIMINARY; PRT; 1865 AA.  
 AC O9C8G1;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Acetyl-CoA carboxylase, putative, 5' partial (Fragment).  
 GN F15C21.1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

ON NCBI\_TaxID=3702;  
[1]  
RX SEQUENCE FROM N.A.  
RC STRAIN=CV\_COLUMBIA,  
RM MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
RA Dunn P., Etgu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
RA Gail J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,  
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,  
RA Miltschew J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,  
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,  
RA Ueberlack T., Van Aken S., Vaysberg M., Vyotskaya V.S., Walker M.,  
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W., Walker M.,  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
RT thaliana."  
RL Nature 408:816-820(2000).  
DR EMBL: AC025781; AACG1250.1; -  
DR InterPro: IPR003439; ABC\_transport.  
DR InterPro: IPR001882; Biotin\_attach.  
DR InterPro: IPR000089; Biotin\_lipoyl.  
DR InterPro: IPR000022; Carboxyl\_trans.  
DR InterPro: IPR000901; CPSase.  
DR InterPro: IPR002016; Peroxidase.  
DR Pfam: PF02785; Biotin\_carb\_C\_1.  
DR Pfam: PF0364; biotin\_lipoyl; 1.  
DR Pfam: PF01039; Carboxyl\_trans; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; UNKNOWN\_1.  
DR PROSITE: PS00188; BIOTIN; UNKNOWN\_1.  
DR PROSITE: PS00436; PEROXIDASE\_2; UNKNOWN\_1.  
FT NON\_TER 1 1  
SQ SEQUENCE 1865 AA; 209161 MW; 9B9FCA50C2846DD0 CRC64;

Query Match 22.2%; Score 63; DB 10; Length 1865;  
Best Local Similarity 27.5%; Pred. NO. 40;  
Matches 14; Conservative 12; Mismatches 23; Indels 2; Gaps 1.

QY 1 YLYOMLGAPVPPDPLEPREVEYCELNPD-CDELADHIGFOEAYRREYGPV 49  
I::I:I:I I:I I:I I:I I:I I:I I:I I:I I:I I:I I:I I:I I:I I:  
Db 1429 YIPAVVGSLPLAPLDPPERIVERPEKSCDPRALAGVKNTGWKGLGI 1479

RESULT 10  
O38970 PRELIMINARY: PRT: 2254 AA.  
AC Q38970  
DT 01-NOV-1996 (TREMBLrel. 01. Created)  
DT 01-NOV-1996 (TREMBLrel. 01. Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21. Last annotation update)  
DE Acetyl-CoA carboxylase.  
GN ACC1.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC euroids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
[1]  
RN RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA ECOTYPE.  
RM MEDLINE=96012930; PubMed=7551584;  
RA Yanai Y., Kawasaki T., Shimada H., Wurtele E.S., Nikolau B.J.,  
RA Ichikawa N.;  
RT "Genomic organization of 251 kDa acetyl-CoA carboxylase genes in  
Arabidopsis: tandem gene duplication has made two differentially  
expressed isozymes.";  
Plant Cell Physiol. 36:779-787(1995).

DR EMBL: D34630; BAA07012.1; -  
DR EMBL: D34631; BAA07013.1; -  
DR HSSP; P24182; IDV1.  
DR InterPro: IPR003439; ABC\_transportr.  
DR InterPro: IPR001882; Biotin\_attach.  
DR InterPro: IPR000089; Biotin\_lipoyl.  
DR InterPro: IPR000022; Carboxyl\_trans.  
DR InterPro: IPR000901; CPSase.  
DR InterPro: IPR002016; peroxidase.  
DR Pfam; PF02785; Biotin\_card\_C; 1.  
DR Pfam; PF00364; biotin\_lipoyl; 1.  
DR Pfam; PF01039; Carboxyl\_trans; 1.  
DR Pfam; PF00289; CPSase\_L\_chain; 1.  
DR Pfam; PF02786; CPSase\_L\_D2; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; UNKNOWN\_1.  
DR PROSITE; PS00188; BIOTIN; UNKNOWN\_1.  
DR PROSITE; PS00867; CPSASE\_2; UNKNOWN\_1.  
DR PROSITE; PS00436; PEROXIDASE\_2; UNKNOWN\_1.  
SO SEQUENCE 2254 AA; 251380 MW; F00A5AE31CC5E87 CRC64;

Query Match 22.2%; Score 63; DB 10; Length 2254;  
Best Local Similarity 27.5%; Pred. No. 49;  
Matches 14; Conservative 12; Mismatches 23; Indels 2; Gaps 1;

DY 1 YLYQWLGAPVPYPDPLEPRREVCENLPD-CDELDADHIGFQEARFRFGPV 49  
Db 1818 YIPAVGGPLVLAPLDPPERIVERYPENSCDPRAIAGVKNDGKWLGI 1868

RESULT 11  
Q38971 PRELIMINARY; PRT; 2254 AA.

AC Q38971;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DI 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
DE Acetyl-CoA carboxylase (Acetyl-CoA carboxylase 1) (EC 6.4.1.2).  
GN ACCL.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
CX NCBI\_TaxID=3702;  
[1]  
RN R1  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RX MEDLINE=94345007; PubMed=7915036;  
RA Roessler K.R., Shorrosh B.S., Ohlrogge J.B.;  
RT "Structure and expression of an Arabidopsis acetyl-coenzyme A  
carboxylase gene.";  
RL Plant Physiol. 105:611-617(1994).  
[2]  
RN R2  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. COLUMBIA;  
RX MEDLINE=96012930; PubMed=7551584;  
RA Yanai Y., Kawasaki T., Shimada H., Wurttele E.S., Nikolau B.J.,  
RA Ichikawa N.;  
RT "Genomic organization of 251 kDa acetyl-CoA carboxylase genes in  
Arabidopsis: tandem gene duplication has made two differentially  
expressed isozymes.";  
RL Plant Cell Physiol. 36:779-787(1995).  
[3]  
RN R3  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. COLUMBIA;  
RA Johnson J.L., Choi J.K., Yanai Y., Wurttele E.S., Nikolau B.J.;  
RT "Structure and expression patterns of the duplicated genes, ACCL and  
thaliana.";  
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.  
EMBL: L27074; AAC41645.1; -  
DR EMBL: AF062308; AAC40563.1; -  
DR HSSP; P24182; IDV1.  
DR InterPro: IPR003439; ABC\_transportr.

ID	Q93NM6	PRELIMINARY:	PRT: 10917 AA.
AC	Q93NM6		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, last annotation update)		
DE	AmphC.		
CN	AMPHC.		
OS	Streptomyces nodosus.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Streptomycetaceae; Streptomyces.		
OX	NCBI_TaxID=40318;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Califery P., Lynch S.V., Flood E.M., Finnian S.M., O'Leary M.;		
RT	"The amphoterich biosynthetic gene cluster from Streptomyces nodosus.";		
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF357202; AAK73514.1; -		
DR	InterPro; IPR001227; AC_transferase.		
DR	InterPro; IPR002328; ADH_zinc.		
DR	InterPro; IPR002085; Adh_zn_family.		
DR	InterPro; IPR004410; Fabd.		
DR	InterPro; IPR000794; ketoacyl-synt.		
DR	InterPro; IPR003880; pantane_attach.		
DR	Pfam; PF00698; Aclyl_transf. 6.		
DR	Pfam; PF00107; adh_zinc 1.		
DR	Pfam; PF00109; ketoacyl-synt. 6.		
DR	Pfam; PF02801; ketoacyl-synt.C; 6.		
DR	Pfam; PF00550; pp-binding. 6.		
DR	TIGRFAMS; TIGR00128; fabd. 6.		
DR	PROSITE; PS50075; ACP_DOMAIN. 6.		
DR	PROSITE; PS00059; ADH_ZINC; UNKNOWN_1.		
DR	PROSITE; PS00606; B_KETOACYL_SYNTHAS; UNKNOWN_6.		
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_5.		
KW	Phosphopantetheine.		
SO	SEQUENCE 10917 AA; 1132905 MW; 15AC5956B5810A1 CRC64;		
Query Match 22.0%; Score 62.5; DB 2; Length 10917;			
Best Local Similarity 41.2%; Pred. No. 3.3e+02;			
Matches 14; Conservative 5; Mismatches 12; Indels 3; Gaps 2;			
QY	4 QMLGAPVPPPLPRRREYCELNPPCCDELADHIG 37		
DB	2937 EW--TPVPLPDG-PTTESIALIGPPTDDLAEMIG 2967		
	: 1 11 11 1 : 1 11 11 : 1		
RESULT 14			
ID	Q09967	PRELIMINARY:	PRT: 574 AA.
AC	Q09967		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)		
DE	Hypothetical 62.2 Kda protein.		
CN	B0244. 8.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;		
OC	Rhabditidae; Pelodermineae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RX	MEDLINE=99069613; PubMed=9851916;		
RA	None;		
RT	"Genome sequence of the nematode C. elegans: a platform for		
RT	investigating biology. The C. elegans Sequencing Consortium.";		
RL	Science 282:2012-2018(1998).		
RN	[2]		

RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Favello A.;  
RT "The sequence of C. elegans cosmid 80244."  
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission."  
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases  
DR EMBL: U28971; AAA8376.1; -  
DR HSSP: P01130; 1LDR  
DR InterPro: IPR002172; LDL\_recept-A.  
DR Pfam: PF00057; 1dl\_recept\_a; 7.  
DR PRINTS; PRO0261; LDLRECEPTOR.  
DR SMART; SM00192; LDLA; 8.  
DR PROSITE; PS01209; LDLRA\_1; 7.  
DR PROSITE; PS50068; LDLRA\_2; 8.  
KW Glycoprotein; Hypothetical protein.  
SQ SEQUENCE 574 AA; 62206 MW; 7405C4AFCD22D239 CRC64;

Query Match	21.8%	Score 62;	DB 5;	Length 574;
Best Local Similarity	61.1%;	Pred. No. 15;		
Matches 11; Conservative	3;	Mismatches 4;	Indels 0;	Gaps 0;

```
QY 19 RREVCENPDCDELADHI 36
    ||::||: ||||: || |
Db 417 RRVCDGTPDCDDGADEI 434
```

RESULT 15				
092KK7				
ID	092KK7	PRELIMINARY;	PRT;	609 AA.
AC	092KK7;			
DT	01-DEC-2001 (TREMblrel, 19, Created)			
DT	01-DEC-2001 (TREMblrel, 19, last sequence update)			
DT	01-MAR-2002 (TREMblrel, 20, last annotation update)			
DE	Hypothetical signal peptide protein SMC00861.			
GN	R00845 OR SMC00861.			
OS	Rhizobium meliloti (Sinorhizobium meliloti).			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group.			
OC	Rhizobiaceae; Sinorhizobium.			
OX	NCBI_TaxID=382;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=1021;			
RA	MEDLINE=21396507; PubMed=11481430;			
RA	Capela D., Barloy-Hubier F., Gonzy J., Bothe G., Ampe F., Batut			
RA	Boisnard P., Becker A., Boutry M., Cadieu E., Dreno S., Gloux S			
RA	Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,			
RA	Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,			
RA	Renard C., Theault P., Vandenbol M., Weidner S., Galbert F.;			
RT	"Analysis of the chromosome sequence of the legume symbiont			
RT	Sinorhizobium meliloti strain 1021."			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:9677-9682(2001).			
DR	EMBL; AL591785; CA045417.1; -.			
DR	InterPro; IPR001440; TPR.			
DR	Pfam; PF00515; TPR. 8			
KO	Hypothetical protein; Complete proteome.			
CO	SEQUENCE 609 AA; 67667 MW; D771FE911E69542 CRC64;			

Query Match	21.7%;	Score 61.5;	DB 16;	Length 609;
Best Local Similarity	37.8%;	Pred. No. 18;		
Matches 17; Conservative	6;	Mismatches 19;	Indels 3;	Gaps 2;

QY 1 YL-YQWLGAVPYPPDPLEPRREVCELNPDCDELADHIGFOEAYRR 44  
|||: |||: |||: |||: |||:  
Db 467 YLGSWVDNINLEDGLEMIKRAVELKPDGGYIVDSLGW--AYFR 509

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:31:34 ; Search time 19 Seconds  
(without alignments)  
247.926 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 284  
Sequence: 1 YLYGNLGAAPPYPPPLPRR.....DELADHIGFORAIFYGPV 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	100.0	100	1	GEHU
2	277	97.5	49	1	osteocalcin precur
3	262	92.3	100	1	osteocalcin - crab
4	237	83.5	49	1	osteocalcin precur
5	220	77.5	99	1	GECT
6	217	76.4	49	1	osteocalcin precur
7	199	70.1	48	2	osteocalcin - rabb
8	186.5	65.7	97	1	osteocalcin - emu
9	166	58.5	95	2	osteocalcin precur
10	162	58.5	95	2	osteocalcin - mus
11	162	57.0	95	2	osteocalcin - mus
12	162	57.0	95	2	osteocalcin - mus
13	109	38.4	47	1	GEWF
14	86	30.3	45	1	osteocalcin - swor
15	63	22.2	2257	1	osteocalcin - blue
16	62.5	21.8	435	2	protein F5j5.19 [i
17	62	22.0	574	2	osteocalcin - blue
18	60.5	21.3	677	1	osteocalcin - relat
19	60.5	21.3	677	2	osteocalcin - relat
20	60.5	21.3	677	2	osteocalcin - relat
21	60	21.1	1344	2	osteocalcin - relat
22	60	21.1	1344	2	osteocalcin - relat
23	59	20.8	262	2	osteocalcin - relat
24	59	20.8	262	2	osteocalcin - relat
25	59	20.8	262	2	osteocalcin - relat
26	58.5	20.6	430	2	osteocalcin - relat
27	58.5	20.6	430	2	osteocalcin - relat
28	58.5	20.6	430	2	osteocalcin - relat
29	58.5	20.6	430	2	osteocalcin - relat

#### ALIGNMENTS

30	57.5	20.2	409	2	T30281	hypothetical prote
31	57	20.1	214	2	T39559	probable ubiquinol
32	57	20.1	265	2	T01187	ribosomal protein
33	57	20.1	376	2	S27976	H+-transporting tw
34	57	20.1	726	1	S73915	virulence-associat
35	57	20.1	1561	2	S46200	acetyl-CoA carboxy
36	56.5	19.9	350	2	B75192	hypothetical prote
37	56.5	19.9	373	2	S32537	erythroid transcri
38	56.5	19.9	431	2	T29850	hypothetical prote
39	56	19.7	265	2	T01203	ribosomal protein
40	56	19.7	377	1	PNWTG	H+-transporting tw
41	56	19.7	415	2	S55617	hypothetical prote
42	56	19.7	416	1	SKX1AG	dermal gland prote
43	56	19.7	536	2	T42606	probable transcrip
44	55.5	19.5	346	2	T45069	8-hydroxy-quinane
45	55.5	19.5	634	2	AC3581	5-dehydro-2-deoxyg

#### RESULT 1

GEHU

osteocalcin precursor [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1980 #sequence\_revision 07-Apr-1994 #text\_change 08-Dec-2000

C:Accession: S12652; C25471; A03301; S08694

R:Kiefer, M.C.; Saphire, A.C.S.; Bauer, D.M.; Barr, P.J.

Nucleic Acids Res. 18, 1909, 1990

A:Title: The cDNA and derived amino acid sequences of human and bovine bone Gla prote

A:Reference number: S12652; MUID:90245603; PMID:2336375

A:Accession: S12652

A:Molecule type: mRNA

A:Residues: 1-100 <KIE>

A:Cross-references: EMBL:X53698; NID:q36092; PIDN:CAA37736.1; PID:q29450

R:Celeste, A.J.; Rosen, V.; Buecker, J.L.; Kitz, R.; Wang, E.A.; Wozney, J.M.

J. Biol. Chem. 255, 8685-8691, 1980

A:Title: Isolation and sequence of the vitamin K-dependent protein from human bone. U

A:Reference number: A91045; MUID:87004555; PMID:3019668

A:Accession: C25471

A:Molecule type: DNA

A:Residues: 1-32,35-100 <CEI>

A:Cross-references: EMBL:X04143; NID:q29449; PIDN:CAA27763.1; PID:q29450

R:Poser, J.W.; Esch, F.S.; Ling, N.C.; Price, P.A.

J. Biol. Chem. 255, 8685-8691, 1980

A:Title: Isolation and sequence of the vitamin K-dependent protein from human bone. U

A:Reference number: A44566; MUID:92222128; PMID:1807167

A:Accession: A03301

A:Molecule type: protein

A:Residues: 52-100 <POS>

R:Calins, J.R.; Williamson, M.K.; Price, P.A.

Anal. Biochem. 199, 93-97, 1991

A:Title: Direct identification of gamma-carboxyglutamic acid in the sequencing of vit

A:Reference number: A44566; MUID:92222128; PMID:1807167

A:Accession: A03301

A:Molecule type: protein

A:Residues: 52-100 <POS>

R:Calins, J.R.; Williamson, M.K.; Price, P.A.

Anal. Biochem. 199, 93-97, 1991

A:Title: Direct identification of gamma-carboxyglutamic acid in the sequencing of vit

A:Reference number: A44566; MUID:92222128; PMID:1807167

A:Accession: A03301

A:Molecule type: protein

A:Residues: 52-100 <POS>

Query Match 100.0%; Score 284; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 6, 5e-27;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPYPDPLEPRREVCELNPPCDELADHIGFOEAYRRFGV 49  
 |||||  
 DB 52 YLYQWLGAPVPYPDPLEPRREVCELNPPCDELADHIGFOEAYRRFGV 100

## RESULT 2

## GEMKI

osteocalcin - crab-eating macaque  
 N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Macaca fascicularis (crab-eating macaque)  
 C>Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 06-Sep-1996  
 C:Accession: A03302  
 R:Hauschka, P. V.; Carr, S. A.; Blemann, K.  
 J. Biochemistry 21, 638-642, 1982  
 A:Title: Primary structure of monkey osteocalcin.  
 A:Reference number: A03302; MUID:82182842; PMID:6978733  
 A:Accession: A03302  
 A:Molecule type: protein  
 A:Residues: 1-49 <HAI>  
 C:Comment: This protein, isolated from bone, binds strongly to apatite.  
 C:Superfamily: osteocalcin  
 C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
 F:9/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:17,21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:23-29/Disulfide bonds: #status experimental

Query Match 97.5%; Score 277; DB 1; Length 49;  
 Best Local Similarity 95.9%; Pred. No. 2, 1e-26;  
 Matches 47; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPYPDPLEPRREVCELNPPCDELADHIGFOEAYRRFGV 49  
 |||||  
 DB 1 YLYQWLGAPVPYPDPLEPRREVCELNPPCDELADHIGFOEAYRRFGV 49

## RESULT 3

## GEBO

osteocalcin precursor - bovine  
 N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 24-Apr-1984 #sequence\_revision 07-Apr-1994 #text\_change 22-Jun-1999  
 C:Accession: S12653; A03303; S08693  
 R:Kiefer, M. C.; Saphire, A. C. S.; Bauer, D. M.; Barr, P. J.  
 Nucleic Acids Res. 18, 1909, 1990  
 A:Title: The cDNA and derived amino acid sequences of human and bovine bone gla protein.  
 A:Reference number: S12652; MUID:90245603; PMID:2336375  
 A:Accession: S12653  
 A:Molecule type: mRNA  
 A:Residues: 1-100 <RIE>  
 A:Cross-references: EMBL:X53699; NID:9719; PIDN:CAA37737.1; PID:9720  
 A:Note: alternative splicing may produce a sequence lacking residues 33-34  
 R:Price, P. A.; Poser, J. W.; Raman, N.  
 Proc. Natl. Acad. Sci. U.S.A. 73, 3374-3375, 1976  
 A:Title: Primary structure of the gamma-carboxyglutamic acid-containing protein from bovine bone.  
 A:Reference number: A03303; MUID:77036749; PMID:1068450  
 A:Accession: A03303  
 A:Molecule type: protein  
 A:Residues: 52-100 <PRI>  
 C:Comment: This protein, isolated from bone, binds strongly to apatite.  
 C:Superfamily: osteocalcin  
 C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
 F:1,51/Domin: signal sequence #status predicted <SIG>  
 F:52-100/Product: osteocalcin #status experimental <MAT>  
 F:60/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:66,72,75/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:74-80/Disulfide bonds: #status experimental

Query Match 92.3%; Score 262; DB 1; Length 100;  
 Best Local Similarity 91.8%; Pred. No. 2, 8e-24;

Matches 45; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPYPDPLEPRREVCELNPPCDELADHIGFOEAYRRFGV 49  
 |||||  
 DB 52 YLYQWLGAPVPYPDPLEPRREVCELNPPCDELADHIGFOEAYRRFGV 100

## RESULT 4

## GECT

osteocalcin - cat  
 N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Felis silvestris catus (domestic cat)  
 C>Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 06-Sep-1996  
 C:Accession: A03304  
 R:Shimomura, H.; Kanai, Y.; Sanada, K.  
 J. Biochem. 96, 405-411, 1984  
 A:Title: Primary structure of cat osteocalcin.  
 A:Reference number: A03304; MUID:85054706; PMID:6334077  
 A:Accession: A03304  
 A:Molecule type: protein  
 A:Residues: 1-49 <SHI>  
 C:Superfamily: osteocalcin  
 C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
 F:9/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:17,21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:23-29/Disulfide bonds: #status experimental

Query Match 83.5%; Score 237; DB 1; Length 49;  
 Best Local Similarity 83.7%; Pred. No. 1, 3e-21;  
 Matches 41; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPYPDPLEPRREVCELNPPCDELADHIGFOEAYRRFGV 49  
 |||||  
 DB 1 YLYQWLGAPVPYPDPLEPRREVCELNPPCDELADHIGFOEAYRRFGV 49

## RESULT 5

## GERT

osteocalcin precursor - rat  
 N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 22-Jun-1999  
 C:Accession: A31856; A31419; A32324; A25471  
 R:Yoon, K.; Rutledge, S. J. C.; Buenaaga, R. F.; Rodan, G. A.  
 Biochemistry 27, 8521-8526, 1988  
 A:Title: Characterization of the rat osteocalcin gene: stimulation of promoter activity by 1,25-(OH)<sub>2</sub>D<sub>3</sub>.  
 A:Reference number: A31856; MUID:89118266; PMID:3365336  
 A:Accession: A31856  
 A:Molecule type: DNA  
 A:Residues: 1-99 <YOC>  
 A:Cross-references: GB:M23637; NID:9340986; PIDN:AAA41761.1; PID:9514962  
 R:Theofan, G.; Haberstroh, L. M.; Price, P. A.  
 DNA 8, 213-221, 1989  
 A:Title: Molecular structure of the rat bone gla protein gene and identification of promoter elements.  
 A:Reference number: A31419; MUID:89231082; PMID:2785907  
 A:Accession: A31419  
 A:Molecule type: DNA  
 A:Residues: 1-99 <THE>  
 A:Cross-references: GB:M25490; NID:9576530; PIDN:AAA53280.1; PID:9576531  
 R:Lian, J.; Stewart, C.; Puchacz, E.; Mackowiak, S.; Shalhoub, V.; Collart, D.; Zambelli, N. J.; Acad. Sci. U.S.A. 86, 1143-1147, 1989  
 A:Title: Structure of the rat osteocalcin gene and regulation of vitamin D-dependent expression.  
 A:Reference number: A32324; MUID:89145200; PMID:2784002  
 A:Accession: A32324  
 A:Molecule type: DNA  
 A:Residues: 1-99 <LIA>  
 A:Cross-references: GB:U04500; NID:9205863; PIDN:AAA41764.1; PID:9205864  
 R:Pan, L. C.; Price, P. A.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 6109-6113, 1985  
 A:Title: The promoter of the rat bone gla protein gene shares homology with the promoter of the rat osteocalcin gene.  
 A:Reference number: A25167; MUID:85296305; PMID:3875856  
 A:Accession: A25167  
 A:Molecule type: mRNA



Query Match	77.5%;	Score 220;	DB 1;	Length 99;
Best Local Similarity	80.9%;	Pred. No. 2.9e-19;		
Matches 38;	Conservative 3;	Mismatches 6;	Indels 0;	Gaps 0;

RESULT 6  
A61280

Query Match	76.48;	Score 217;	DB 1;	Length 49;
Best Local Similarity	83.78;	Pred. No. 3.1e-19;		
Matches	36;	Conservative	4;	Mismatches 3;
			Indels	0;
			Gaps	0;

RESULT 7  
S02208

R; Huq, N.L.; Tseng, A.; Chapman Blochem. Int. 15, 271-277, 1987

A;Molecule type: protein  
A;Residues: 1-48 <HUG>

Query Match	70.1%;	Score 199;	DB 2;	Length 48;
Best Local Similarity	94.6%;	Pred. NO. 4.3e-17;		
Matches 35; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0;

RESULT 8  
GECH

C;Species: Gallus gallus (chicken)

R;Neugebauer, B.M.; Moore, J. Bone Miner. Res. 10, 157

A; Accession: I50700

A;Residues: 1-97 <NEU>  
A;Cross-references: EN

A;Title: Gas chromatographic mass spectrometry

A;Molecule type: protein  
A;Residues: 49-63, 'I', 65

C;Keywords: bone; calcium

F:71-77/Disulfide bonds: #status predicted

[illegible]

B25471  
osteocalcin precursor - mouse

C;Accession: B25471; A49871; I611899871;

A;Title: Isolation of the human gene for bone gla protease  
A;Reference number: A91045; MUID:87004555; PMID:30196688

A/Residues: 1-95 <CEL>

A;Title: The mouse osteocalcin gene cluster contains three genes with two separate sp  
A;Reference number: A49871; MUID:94117426; PMID:8288580

[illegible]

A/Residues: L-95 <RES>  
A/Cross-references: GB:L24429; NID:g455452; PIDN:AAA39854.1; PID:g455453

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-3, '1F', 6-10, 'T', 12-95 <RE2>  
A:Cross-references: GB:I24431; NID:945545456; PIDN:AAA39856.1; PID:945545457  
C:Genetics:  
A:Introns: 22/1, 33/1, 52/2  
C:Superfamily: osteocalcin  
F:1-49/Domain: signal sequence #status predicted <SIG>  
F:50-95/Product: osteocalcin #status predicted <MAT>

Query Match 58.5%; Score 166; DB 2; Length 95;  
Best Local Similarity 65.2%; Pred. No. 7, 8e-13;  
Matches 30; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 LYQWLGAPVPPDLPPEPRREVCENLPDDELADHIGFOEAYRRFYG 47  
| : : : : | | | | | | | | | | | | | | | | : : : : |  
Db 47 LRRYLGASVSPDLPPEPRREVCENLPDDELADHIGFOEAYRRFYG 92

RESULT 10  
153275  
osteocalcin - mouse

C:Species: Mus sp. (mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 20-Aug-1999

C:Accession: I53275

R:Rahman, S.; Oberdorf, A.; Montecino, M.; Tanhauser, S.M.; Lian, J.B.; Stein, G.S.; Lai  
Endocrinology 133, 3050-3053, 1993

A:Title: Multiple copies of the bone-specific osteocalcin gene in mouse and rat.

A:Reference number: I53275; MUID:94062692; PMID:8243336

A:Accession: I53275

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: GB:S67455; NID:9456854; PIDN:AAB29145.1; PID:9456856

C:Genetics:  
A:Introns: 22/1, 33/1, 52/2, 72/2

C:Superfamily: osteocalcin

Query Match 58.5%; Score 166; DB 2; Length 95;  
Best Local Similarity 65.2%; Pred. No. 7, 8e-13;  
Matches 30; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 LYQWLGAPVPPDLPPEPRREVCENLPDDELADHIGFOEAYRRFYG 47  
| : : : : | | | | | | | | | | | | | | | | : : : : |  
Db 47 LRRYLGASVSPDLPPEPRREVCENLPDDELADHIGFOEAYRRFYG 92

RESULT 11  
167413  
osteocalcin - mouse

C:Species: Mus sp. (mouse)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Aug-1999

C:Accession: I67413

R:Rahman, S.; Oberdorf, A.; Montecino, M.; Tanhauser, S.M.; Lian, J.B.; Stein, G.S.; Lai

Endocrinology 133, 3050-3053, 1993

A:Title: Multiple copies of the bone-specific osteocalcin gene in mouse and rat.

A:Reference number: I53275; MUID:94062692; PMID:8243336

A:Accession: I67413

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: GB:S67456; NID:9456855; PIDN:AAB29146.1; PID:9456857

C:Genetics:  
A:Introns: 22/1, 33/1, 52/2, 72/2

C:Superfamily: osteocalcin

Query Match 57.0%; Score 162; DB 2; Length 95;  
Best Local Similarity 63.0%; Pred. No. 2, 3e-12;  
Matches 29; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 2 LYQWLGAPVPPDLPPEPRREVCENLPDDELADHIGFOEAYRRFYG 47  
| : : : : | | | | | | | | | | | | | | | | : : : : |  
Db 47 LRRYLGASVSPDLPPEPRREVCENLPDDELADHIGFOEAYRRFYG 92

RESULT 12  
161188  
osteocalcin-related protein - mouse

C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 20-Aug-1999

C:Accession: I61188; I49073

R:Desbols, C.; Hogue, D.A.; Karsenty, G.

J. Biol. Chem. 269, 1183-1190, 1994

A:Title: The mouse osteocalcin gene cluster contains three genes with two separate sp

A:Reference number: A49871; MUID:94117426; PMID:8288580

A:Accession: I61188

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: EMBL:U11541; NID:9508297; PIDN:AAB60445.1; PID:9508298

A:Accession: I49073

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: EMBL:U11541; NID:9508297; PIDN:AAB60445.1; PID:9508298

C:Genetics:  
A:Introns: 22/1, 33/1, 52/2

C:Superfamily: osteocalcin

Query Match 57.0%; Score 162; DB 2; Length 95;  
Best Local Similarity 63.0%; Pred. No. 2, 3e-12;  
Matches 29; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 2 LYQWLGAPVPPDLPPEPRREVCENLPDDELADHIGFOEAYRRFYG 47  
| : : : : | | | | | | | | | | | | | | | | : : : : |  
Db 47 LRRYLGASVSPDLPPEPRREVCENLPDDELADHIGFOEAYRRFYG 92

RESULT 13  
GEMF  
osteocalcin - swordfish

N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing prote

C:Species: Xiphias gladius (swordfish)

C:Date: 30-Apr-1979 #sequence\_revision 27-Nov-1985 #text\_change 06-Sep-1996

C:Accession: A03306

R:Price, P.A.; Otsuka, A.S.; Poser, J.W.

in Calcium-binding proteins and Calcium Function, Wasserman, R.H., Corradino, R.A., C

A:Title: Comparison of gamma-carboxyglutamic acid-containing proteins from bovine and

A:Reference number: A03306

A:Accession: A03306

A:Molecule type: protein

A:Residues: 1-47 <PRI>

A:Note: residues 14, 24, and 37 were not positively identified

C:Superfamily: osteocalcin

C:Keywords: bone; calcium binding; carboxyglutamic acid

F:13,17,20/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental

F:19-25/Disulfide bonds: #status predicted

Query Match 38.4%; Score 109; DB 1; Length 47;  
Best Local Similarity 56.8%; Pred. No. 2, 4e-06;  
Matches 21; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 13 PDLPEPRREVCENLPDDELADHIGFOEAYRRFYGV 49  
| | | | | | | | | | | | | | | | : : : : |  
Db 9 PLQLESLREVCENLNSCDMDATGAIYAAYVYGP 45

RESULT 14  
A42794  
osteocalcin - bluegill

N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing prote

C:Species: Lepomis macrochirus (bluegill)

C:Date: 31-Dec-1993 #sequence\_revision 06-Sep-1996 #text\_change 06-Sep-1996

C:Accession: A42794

R:Nishimoto, S.K.; Araki, N.; Robinson, F.D.; Waite, J.H.

J. Biol. Chem. 267, 11600-11605, 1992





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:33:39 ; Search time 12 Seconds

(without alignments)  
250,138 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 284  
Sequence: 1 YLYQWLGAPVPPDPLEPRR.....DELADHIGFOAYRRFGPV 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 193982 seqs, 61258239 residues

Total number of hits satisfying chosen parameters: 193982

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing First 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	100.0	100	6	US-10-283-656-1
2	242	85.2	58	5	US-09-724-676-90318
3	242	85.2	58	5	US-09-724-676-90318
4	76	26.8	13	5	US-09-802-154-21
5	76	26.8	257	6	US-10-096-246-2
6	57	20.1	314	5	US-09-724-676-48669
7	57	20.1	314	5	US-09-724-676-48669
8	55	19.4	117	1	PCT-US02-32727-3145
9	54	19.0	92	1	PCT-US02-32727-11781
10	53.5	18.8	194	5	US-09-724-676-93282
11	53.5	18.8	194	5	US-09-724-676-93283
12	53.5	18.8	194	5	US-09-724-676-93284
13	53.5	18.8	194	5	US-09-724-676-93285
14	53.5	18.8	194	5	US-09-724-676-93286
15	53.5	18.8	194	5	US-09-724-676-93287
16	53.5	18.8	194	5	US-09-724-676-93288
17	53.5	18.8	194	5	US-09-724-676-93289
18	53.5	18.8	194	5	US-09-724-676-93290
19	53.5	18.8	194	5	US-09-724-676-93291
20	53.5	18.8	194	5	US-09-724-676-93292
21	53.5	18.8	194	5	US-09-724-676-93293
22	53.5	18.8	194	5	US-09-724-676-93294
23	53.5	18.8	194	5	US-09-724-676-93295
24	53.5	18.8	194	5	US-09-724-676-93296
25	53.5	18.8	194	5	US-09-724-676-93297
26	53.5	18.8	194	5	US-09-724-676-93298

#### ALIGNMENTS

27	53.5	18.8	194	5	US-09-724-676-93301	Sequence 93301, A
28	53.5	18.8	194	5	US-09-724-676-93282	Sequence 93282, A
29	53.5	18.8	194	5	US-09-724-676-93283	Sequence 93283, A
30	53.5	18.8	194	5	US-09-724-676-93284	Sequence 93284, A
31	53.5	18.8	194	5	US-09-724-676-93285	Sequence 93285, A
32	53.5	18.8	194	5	US-09-724-676-93286	Sequence 93286, A
33	53.5	18.8	194	5	US-09-724-676-93287	Sequence 93287, A
34	53.5	18.8	194	5	US-09-724-676-93288	Sequence 93288, A
35	53.5	18.8	194	5	US-09-724-676-93289	Sequence 93289, A
36	53.5	18.8	194	5	US-09-724-676-93290	Sequence 93290, A
37	53.5	18.8	194	5	US-09-724-676-93291	Sequence 93291, A
38	53.5	18.8	194	5	US-09-724-676-93292	Sequence 93292, A
39	53.5	18.8	194	5	US-09-724-676-93293	Sequence 93293, A
40	53.5	18.8	194	5	US-09-724-676-93294	Sequence 93294, A
41	53.5	18.8	194	5	US-09-724-676-93295	Sequence 93295, A
42	53.5	18.8	194	5	US-09-724-676-93296	Sequence 93296, A
43	53.5	18.8	194	5	US-09-724-676-93297	Sequence 93297, A
44	53.5	18.8	194	5	US-09-724-676-93298	Sequence 93298, A
45	53.5	18.8	194	5	US-09-724-676-93299	Sequence 93299, A

RESULT 1  
US-10-283-656-1  
; Sequence 1, Application US/10283656  
; GENERAL INFORMATION:  
; APPLICANT: EKEMA, George Mbella  
; APPLICANT: MAYS, Robert W.  
; APPLICANT: BRUNDEN, Kurt R.  
; TITLE OF INVENTION: Methods for Using Osteocalcin  
; FILE REFERENCE: ATX-005  
; CURRENT APPLICATION NUMBER: US/10/283, 656  
; CURRENT FILING DATE: 2002-10-29  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-283-656-1

Query Match 100.0%; Score 284; DB 6; Length 100;  
Best Local Similarity 100.0%; Pred. No. 5.6e-27;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLEPRRVCCLNPDDELADHIGFOAYRRFGPV 49  
Db 52 YLYQWLGAPVPPDPLEPRRVCCLNPDDELADHIGFOAYRRFGPV 100

RESULT 2  
US-09-724-676-90318  
; Sequence 90318, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: CompuGen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 CompuGen  
; CURRENT APPLICATION NUMBER: US/09/724, 676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 90318  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-724-676-90318

Query Match 85.2%; Score 242; DB 5; Length 58;  
Best Local Similarity 97.7%; Pred. No. 3.4e-22;  
Matches 43; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



```

; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siding
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 3145
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Propionl acnes
PCT-US02-32727-3145
```

```

Query Match          19.4%; Score 55; DB 1; Length 117;
Best Local Similarity 30.8%; Pred. No. 14;
Matches 12; Conservative 7; Mismatches 8; Indels 12; Gaps 2;
```

```

OY      3 YWLGAPVY-----PDPLE--PREVCELNPDC 29
Db      12 YWVWGLGPPINDPPIRAPDPDAVSTPKAVCKESSQC 50
```

```

RESULT 9
PCT-US02-32727-11781
; Sequence 11781, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siding
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 11781
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Propionl acnes
PCT-US02-32727-11781
```

```

Query Match          19.0%; Score 54; DB 1; Length 92;
Best Local Similarity 31.2%; Pred. No. 15;
Matches 15; Conservative 5; Mismatches 28; Indels 0; Gaps 0;
```

```

OY      1 YLYWLGAPVYPPDPLEPREVCELNPDCDELADHIGFOEAYRRFGP 48
Db      15 FLVWMLVADILRLSPVTPRLIVSMASPTPSKCRTWCSMAAVRPPVP 62
```

```

RESULT 10
US-09-724-676-93282
; Sequence 93282, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 93282
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-93282
```

```

Query Match          18.8%; Score 53.5; DB 5; Length 194;
Best Local Similarity 36.6%; Pred. No. 34;
Matches 15; Conservative 1; Mismatches 12; Indels 13; Gaps 3;
```

```

OY      8 APVYPPDPLEPREVCELNPDC-----ELADHIG 37
Db      93 APVP-PPPALPRRFICSF-PDCSANYSKAWKLDADHLCRHTG 131
```

```

RESULT 11
US-09-724-676-93283
; Sequence 93283, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 93283
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-93283
```

```

Query Match          18.8%; Score 53.5; DB 5; Length 194;
Best Local Similarity 36.6%; Pred. No. 34;
Matches 15; Conservative 1; Mismatches 12; Indels 13; Gaps 3;
```

```

OY      8 APVYPPDPLEPREVCELNPDC-----ELADHIG 37
Db      93 APVP-PPPALPRRFICSF-PDCSANYSKAWKLDADHLCRHTG 131
```

```

RESULT 12
US-09-724-676-93284
; Sequence 93284, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 93284
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-93284
```

```

Query Match          18.8%; Score 53.5; DB 5; Length 194;
Best Local Similarity 36.6%; Pred. No. 34;
Matches 15; Conservative 1; Mismatches 12; Indels 13; Gaps 3;
```

QY 8 APVPYDPLERREVELNPDCD-----ELADHG 37  
 ||| | | | | : | | |  
 Db 93 APTP-DRAPEPRRLCSF-PDCSANYSKAMKLDALHCKHTG 133

RESULT 13  
US-09-724-676-93285  
; Sequence 93285, Application US/09724676

```

1 APPLICANT: CompuGen LTD
2
3 TITLE OF INVENTION: Variants of alternative splicing
4
5 FILE REFERENCE: 129181.4 CompuGen
6
7 CURRENT APPLICATION NUMBER: US/09/724,676
8
9 CURRENT FILING DATE: 2000-11-28
10
11 NUMBER OF SEQ. ID NOS.: 97222
12
13 SOFTWARE: PatentIn version 3.2
14
15 SEQ. ID NO. 01266

```

ORGANISM: Homo sapiens  
US-09-724-676-93285

Query Match	18.8%	Score 53.5	DB 5	Length 194
Best Local Similarity	36.6%	Pred. No. 34		
Matches 15, Conservative	1	Mismatches 12	Indels 13	Gaps 3

QY 8 APVPYDPLERREVELNPDCD-----ELADHG 37  
 ||| ||| : |||  
 Db 93 APTP-DRAPLRRFICSF-PDCSANYSKAMKLDHICKHG 133

RESULT 14  
US-09-724-676-93286

1. APPLICANT: Comogen Ltd  
 2. TITLE OF INVENTION: Variants of alternative splicing  
 3. FILE REFERENCE: 129181, 4 Comogen  
 4. CURRENT APPLICATION NUMBER: US/09/724, 676  
 5. CURRENT FILING DATE: 2000-11-28  
 6. NUMBER OF SEQ ID NOS: 97222  
 7. SOFTWARE: PatentIn version 3.2

ORGANISM: Homo sapiens  
US-09-724-676-93286

Query Match	18.8%	Score	53.5	DB	5	Length	194
Best Local Similarity	35.6%	Pred. No.	34				
Matches	15	Conservative	1	Mismatches	12	Indels	13
						Gaps	3

```

07      8  APVPIPDPLEPRRVEGELNPDCCD-----ELADHG 37
      ||| ||| : |||
Db    93  APTP-PRPALPRRFISF-PPCSANYSKAKLDAHLCKHTG 133

```

RESULT 15  
US-09-724-676-93287

APPLICANT: Comugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Comugen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: Patent version 3.2

ORGANISM: Homo sapiens  
JUS-09-724-676-93287

Query Match	18.8%;	Score 53.5;	DB 5;	Length 194;
Best Local Similarity	36.6%;	Pred. No. 34;		
Matches 15; Conservative	1;	Mismatches 12;	Indels 13;	Gaps 3;

```

QY      8  APVPYPPDPLEPRREVCNELNPDCD-----ELADHIG 37
      ||| ||| : ||| |||
DB     93  APTP-PRPALPRREICSF-PDCSANYSKAWKLDALHLCCKHTG 131

```

```
Search completed: December 4, 2002, 15:37:41
Job time : 12 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 14:55:59 ; Search time 11 Seconds

(without alignments)  
184.758 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 284

Sequence: 1 YLYQWLGAPVPPDLEPRR.....DELADHIGFQEARFRGPGV 49

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	284	100.0	100	1 OSTC_HUMAN	P02818 homo sapien
2	277	97.5	49	1 OSTC_MACRA	P02819 macaca fasc
3	265	93.3	49	1 OSTC_HORSE	P83005 equus caball
4	262	92.3	100	1 OSTC_BOVIN	P02820 bos taurus
5	244	85.9	49	1 OSTC_CANRA	P81455 canis famill
6	237	83.5	49	1 OSTC_FELCA	P02821 felis silve
7	220	77.5	99	1 OSTC_RAT	P04640 rattus norv
8	217	76.4	49	1 OSTC_RABIT	P35056 oryctolagus
9	199	70.1	48	1 OSTC_DRONO	P15504 diromais no
10	186.5	65.7	97	1 OSTC_CHICK	P02822 gallus gall
11	186	65.5	49	1 OSTC_XENLA	P40147 xenopus lae
12	166	58.5	95	1 OSTC_MOUSE	P04641 mus musculu
13	162	57.0	95	1 OSTC_MOUSE	P54615 mus musculu
14	109	38.4	47	1 OSTC_XIPGL	P02823 xiphias gla
15	89	31.3	45	1 OSTC_SPAUV	P02824 sparus auct
16	86	30.3	45	1 OSTC_LEPMA	P28317 lepomis mac
17	73	25.7	102	1 MGP_GALGA	P36820 galeorhinus
18	68	23.9	103	1 MGP_CHICK	P42413 gallus gall
19	62.5	22.0	435	1 DCOR_PANRE	P49725 panagrellus
20	60.5	21.3	676	1 SYM_ECOLI	P00959 escherichia
21	60	21.1	393	1 ACRC_RHIME	O94449 trichobium m
22	59.5	21.0	203	1 PAX1_CHICK	P47236 gallus gall
23	59	20.8	262	1 R54_GOSHI	P46299 gossypium h
24	59	20.8	264	1 R54_SOLTU	P46300 solanum tub
25	59	20.6	384	1 SX18_HUMAN	P35713 homo sapien
26	58.5	20.6	430	1 Y588_METJA	O58008 methanococc
27	58	20.4	261	1 GUN4_TRIRE	O14405 trichoderna
28	57	20.1	261	1 R54_PPRAR	O81363 prunus arme
29	57	20.1	265	1 R54_MAIZE	O22424 zea mays (m
30	57	20.1	376	1 ATRG_PEA	P28552 pisum sativ
31	57	20.1	726	1 RNR_MYCPN	P75529 mycoplasma
32	56	19.7	225	1 TN12_MOUSE	O54907 mus musculu
33	56	19.7	377	1 ATRPG_TOBAC	P29790 nicotiana t

34	56	19.7	377	1 SX18_MOUSE	P43680 mus musculu
35	56	19.7	439	1 XP2_XENLA	P17437 xenopus lae
36	55.5	19.5	345	1 XPG1_HUMAN	O15527 homo sapien
37	55	19.4	210	1 MODA_BRASN	P50347 bradyrhizob
38	55	19.4	364	1 ATRG_SPIOL	P05435 spinacia ol
39	55	19.4	373	1 ATRP_ARATH	O01908 arabidopsis
40	55	19.4	447	1 VANS_ENTFA	O47745 entercoccu
41	55	19.4	759	1 Z287_MOUSE	O946b9 mus musculu
42	55	19.4	1327	1 YNCA_CAEEL	P34544 caenorhabd1
43	55	19.4	2326	1 PG52_RAT	O00657 rattus norv
44	54.5	19.2	486	1 ENY_HTLV2	P03383 human t-cel
45	54.5	19.2	828	1 YKR6_YEAST	P34239 saccharomyc

## ALIGNMENTS

## RESULT 1

ID	OSTC_HUMAN	STANDARD:	PRT:	100 AA.
AC	P02818:			
DT	21-JUL-1986 (rel. 01, Created)			
DT	01-AUG-1990 (rel. 15, Last sequence update)			
DT	15-JUN-2002 (rel. 41, Last annotation update)			
DE	Osteocalcin precursor (gamma-carboxyglutamic acid-containing protein)			
DE	(Bone Gla-protein) (BCP).			
GN	BGLAP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90245603; PubMed=2336375;			
RA	Kiefer M.C., Saphire A.C.S., Bauer D.M., Barr P.J.;			
RT	"The cDNA and derived amino acid sequences of human and bovine bone			
RT	Gla protein.";			
RL	Nucleic Acids Res. 18:1909-1909(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87004555; PubMed=3019668;			
RA	Celaste A.J., Buecker J.L., Kriz R., Wang E.A., Mooney J.M.;			
RT	"Isolation of the human gene for bone gla protein utilizing mouse and			
RT	rat cDNA clones.";			
RL	EMBO J. 5:1885-1890(1986).			
RN	[3]			
RP	SEQUENCE OF 52-100.			
RX	MEDLINE=81006914; PubMed=6967872;			
RA	Poser J.W., Esch F.S., Ling N.C., Price P.A.;			
RT	"Isolation and sequence of the vitamin K-dependent protein from human			
RT	bone. Undercarboxylation of the first glutamic acid residue.";			
RL	J. Biol. Chem. 265:8685-8691(1980).			
CC	-1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS			
CC	STRONGLY TO APARTITE AND CALCIUM.			
CC	-1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K			
CC	DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE			
CC	BINDING OF CALCIUM.			
CC	-1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN			
CC	FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
DR	EMBL; X53688; CAA37736.1; -			
DR	EMBL; X51699; CAA35996.1; -			
DR	EMBL; X04143; CAA27763.1; -			
DR	PIR; A03301; GEHU.			
DR	PIR; C25471; C25471.			

DR PIR: S12652; S12652.  
DR Genew: HGNC:1043; BGLAP.  
DR MIM: 112260; -;  
DR InterPro: IPR002384; GLA\_bone.  
DR InterPro: IPR00294; VitK\_dep\_GLA.  
DR Pfam: PF00594; gla; 1.  
DR PRINTS: PR00002; GLABONE.  
DR SMART: SM00069; GLA; 1.  
DR PROSITE: PS00011; GLU CARBOXYLATION; 1.  
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Bone; Signal.  
FT SIGNAL 1 23 PROBABLE.  
FT PROPEP 24 51 PROBABLE.  
FT CHAIN 52 100 OSTEOCALCIN.  
FT MOD\_RES 68 68 GAMMA-CARBOXYGLUTAMIC ACID (IN 9% OF THE  
MOLECULES).  
FT MOD\_RES 72 72 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 75 75 GAMMA-CARBOXYGLUTAMIC ACID.  
FT DISULFID 74 80  
FT CONFLICT 33 34 MISSING (IN REF. 2).  
SQ SEQUENCE 100 AA; 10962 MW; 4DF2A0A80849CB71 CRC64;  
Query Match 100.0%; Score 284; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 7.4e-29;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 YLYQWLGAAPVPPDLEPRREVCLEINPDCDELADHIGFQEAAYRRFGYV 49  
|||||  
Db 52 YLYQWLGAAPVPPDLEPRREVCLEINPDCDELADHIGFQEAAYRRFGYV 100  
RESULT 2  
OSTC\_MACFA STANDARD; PRT; 49 AA.  
ID OSTC\_MACFA  
AC P02819;  
DT 21-JUN-1986 (Rel. 01, Created)  
DT 21-JUN-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-  
protein) (BGP).  
GN BGLAP.  
OS Macaca fascicularis (Crah eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=82182842; PubMed=6978733;  
RA Hauschka P.V., Carr S.A., Blemann K.;  
RT "Primary structure of monkey osteocalcin."  
RL Biochemistry 21:638-642(1982).  
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
STONGLY TO APATITE AND CALCIUM.  
CC -1- PWM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
BINDING OF CALCIUM.  
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
FAMILY.  
DR InterPro: IPR002384; GLA\_bone.  
DR InterPro: IPR00294; VitK\_dep\_GLA.  
DR Pfam: PF00594; gla; 1.  
DR PRINTS: PR00002; GLABONE.  
DR SMART: SM00069; GLA; 1.  
DR PROSITE: PS00011; GLU CARBOXYLATION; 1.  
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;  
Bos.  
FT MOD\_RES 9 9 HYDROXYLATION.  
FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
FT DISULFID 23 29 BY SIMILARITY.  
SQ SEQUENCE 49 AA; 5743 MW; C20116014D0C4958 CRC64;

Query Match 97.5%; Score 277; DB 1; Length 49;  
Best Local Similarity 95.9%; Pred. No. 2.5e-28;  
Matches 47; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 YLYQWLGAAPVPPDLEPRREVCLEINPDCDELADHIGFQEAAYRRFGYV 49  
|||||  
Db 1 YLYQWLGAAPVPPDLEPRREVCLEINPDCDELADHIGFQEAAYRRFGYV 49  
RESULT 3  
OSTC\_HORSE STANDARD; PRT; 49 AA.  
ID OSTC\_HORSE  
AC P83005;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone GLA-  
protein) (BGP).  
GN BGLAP.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE, HYDROXYLATION, AND GAMMA-CARBOXYGLUTAMIC ACIDS.  
RC TISSUE=Bone;  
RA Carstensen B., Matilez R., Amory H., Lepeage O.M., Remy B.;  
RT "Isolation and characterization of equine osteocalcin."  
RL Ann. Med. Vet. 0:0-0(2002).  
CC -1- FUNCTION: Constitutes 1-2% of the total bone protein. It binds  
strongly to apatite and calcium.  
CC -1- PWM: Gamma-carboxyglutamic acid residues are formed by vitamin K  
dependent carboxylation. These residues are essential for the  
binding of calcium.  
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
FAMILY.  
DR InterPro: IPR00294; VitK\_dep\_GLA.  
DR Pfam: PF00594; gla; 1.  
DR SMART: SM00069; GLA; 1.  
DR PROSITE: PS00011; GLU CARBOXYLATION; 1.  
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K;  
Hydroxylation.  
FT MOD\_RES 9 9 HYDROXYLATION.  
FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
FT DISULFID 23 29 BY SIMILARITY.  
SQ SEQUENCE 49 AA; 5732 MW; A5B826014D12857F CRC64;  
Query Match 93.3%; Score 265; DB 1; Length 49;  
Best Local Similarity 93.9%; Pred. No. 7.9e-27;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 YLYQWLGAAPVPPDLEPRREVCLEINPDCDELADHIGFQEAAYRRFGYV 49  
|||||  
Db 1 YLYQWLGAAPVPPDLEPRREVCLEINPDCDELADHIGFQEAAYRRFGYV 49  
RESULT 4  
OSTC\_BOVIN STANDARD; PRT; 100 AA.  
ID OSTC\_BOVIN  
AC P02820;  
DT 21-JUN-1986 (Rel. 01, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)  
(Bone Gla-protein) (BGP).  
GN BGLAP.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.

```

OX      NCB1_TaxID=9913;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=90245603; PubMed=2336375;
RX      Kiefer M.C., Saphire A.C.S., Bauer D.M., Barr P.J.;
RT      "The cDNA and derived amino acid sequences of human and bovine bone
RL      gla protein".
RN      Nucleic Acids Res. 18:1909-1909(1990).
RP      [2]
RP      SEQUENCE OF 52-100.
RX      MEDLINE=77036749; PubMed=1068450;
RA      Price P.A., Poser J.M., Raman N.;
RT      "Primary structure of the gamma-carboxyglutamic acid-containing
RL      protein from bovine bone.".
CC      Proc. Natl. Acad. Sci. U.S.A. 73:3374-3375(1976).
CC      -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC      STRONGLY TO APATITE AND CALCIUM.
CC      -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC      DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC      BINDING OF CALCIUM.
CC      -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X53699; CAA37737.1; -.
DR      EMBL; X51700; CAA35997.1; -.
DR      PIR; A03303; GEBO.
DR      PIR; S12653; S12653.
DR      InterPro: IPR002384; GLA_Dome.
DR      InterPro: IPR00294; VitK_dep_GLA.
DR      Pfam; PF00594; gla; 1.
DR      PRINTS; PR00002; GLABONE.
DR      SMART; SM00069; GLA; 1.
DR      PROSITE; PS00011; GLU CARBOXYLATION; 1.
KW      Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;
KM      Bone; Signal.
FT      SIGNAL          1      23
FT      PROPEP         24      51      PROBABLE.
FT      CHAIN          52     100      OSTEOCALCIN.
FT      MOD_RES        60      60      HYDROXYLATION.
FT      MOD_RES        68      68      GAMMA-CARBOXYGLUTAMIC ACID.
FT      MOD_RES        72      72      GAMMA-CARBOXYGLUTAMIC ACID.
FT      MOD_RES        75      75      GAMMA-CARBOXYGLUTAMIC ACID.
FT      DISULFID       74      80
SQ      SEQUENCE      100 AA; 11042 MW; 73015D1681B26219 CRC64;
Query Match          92.3%; Score 262; DB 1; Length 100;
Best Local Similarity 91.8%; Pred. No. 4,2e-26;
Matches 45; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY      1 YLYQGLGAPVYPDPLEPRREVCELNPPCDELADHIGFGEAARREYGPV 49
      |||||||
Db      52 YLDHHLGNAPRYPDPLEPRREVCELNPPCDELADHIGFGEAARREYGPV 100
RESULT 5
OSTC_CANFA          STANDARD;          PRT;          49 AA.
AC      P81455;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bgp).
DE      BGLAP.
OS      Canis familiaris (Dog).

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93318657; PubMed=8101026;  
 RA Colombo G., Pantl P., Yao C., Malluche H.H.;  
 RT "Isolation and complete amino acid sequence of osteocalcin from  
 RL canine bone." Res. 8:733-743(1993).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTR: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; vitk\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR000002; GLABONE.  
 DR SMART: SM00069; GLA; 1.  
 DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.  
 KW Calcium-binding; gamma-carboxyglutamic acid; vitamin K; Hydroxylation;  
 KW Bone.  
 KM MOD\_RES 9 9 HYDROXYLATION (BY SIMILARITY).  
 FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 23 29  
 SQ SEQUENCE 49 AA: 5524 MW: 43121D015817CEA6 CRC64;  
  
 Query Match 85.9%; Score 244; DB 1; Length 49;  
 Best Local Similarity 87.8%; Pred. No. 3.4e-24;  
 Matches 43; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
 OY 1 YLYQWLGAPVPPDPLEPREVCELNPDCELADHIGFOEAYRFFGPV 49  
 || |||||  
 DB 1 YDSDLGAPVPPDPLEPREVCELNPDCELADHIGFOEAYQRFPGPV 49  
  
 RESULT 6  
 OSTC\_FELCA STANDARD; PRT; 49 AA.  
 AC P02821;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-  
 DE protein) (BGP).  
 GN BGLP.  
 OS Felis silvestris catus (Cat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=85054706; PubMed=6334077;  
 RA Shimomura H., Kanai Y., Sanada K.;  
 RT "Primary structure of cat osteocalcin." J. Biochem. 96:405-411(1984).  
 RL J. Biochem. 96:405-411(1984).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTR: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 DR PIR: A03304; GECY.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; vitk\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR000002; GLABONE.



```

CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- P-TM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
DR PIR: A61280; A61280.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; Vitr_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;
KM Bone.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 9 9 HYDROXYLATION.
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 23 29 BY SIMILARITY.
SQ SEQUENCE 49 AA; 5431 MM; 7B218871F0312253 CRC64;

Query Match 76.4%; Score 217; DB 1; Length 49;
Best Local Similarity 83.7%; Pred. No. 8, 1e-21;
Matches 36; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 GABVPPDPLEPREVCELPDDELADHIGFQEAAYRRFGPV 49
DB 7 GABVPPDPLEPREVCELPDDELADHIGFQEAAYRRFGPV 49

RESULT 9
OSTC_DRONO
ID OSTC_DRONO STANDARD: PRT: 48 AA.
AC P15504;
DR 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
OS BGLAP.
OS Dromaius novaehollandiae (Emu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Casuariiformes; Dromaiidae;
OC Dromaius.
ON NCBI_TaxID-8790;
RN [1]
RP SEQUENCE.
RX MEDLINE-88134266; PubMed-3501719;
RA Hug N.L., Tseng A., Chapman G.E.;
RT The amino acid sequence of Emu osteocalcin: gas phase sequencing of
RT Gla-containing proteins.
RL Biochem. Int. 15:271-277(1987).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- P-TM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
DR PIR: S02208; S02208.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; Vitr_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Bone.
FT MOD_RES 16 16 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 20 20 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 23 23 GAMMA-CARBOXYGLUTAMIC ACID.

```

```

FT DISULFID 22 28 BY SIMILARITY.
SQ SEQUENCE 48 AA; 5292 MM; 50A00F3BFABC7FFD CRC64;

Query Match 70.1%; Score 199; DB 1; Length 48;
Best Local Similarity 94.6%; Pred. No. 1, 4e-18;
Matches 35; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 PDPLEPREVCELPDDELADHIGFQEAAYRRFGPV 49
DB 12 PDPLEPREVCELPDDELADHIGFQEAAYRRFGPV 48

RESULT 10
OSTC_CHICK
ID OSTC_CHICK STANDARD: PRT: 97 AA.
AC P02822; Q90620;
DR 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin precursor (gamma-carboxyglutamic acid-containing protein)
DE (bone gla-protein) (BGP).
OS Gallus gallus (Chicken).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
ON NCBI_TaxID-9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-White leghorn; TISSUE-Bone;
RX MEDLINE-95266465; PubMed-7747623;
RA Neugebauer B.M., Moore M.A., Broess M., Gerstenfeld L.C.,
RA Hauschka P.V.;
RT Characterization of structural sequences in the chicken osteocalcin
RT gene: expression of osteocalcin by maturing osteoblasts and by
RT hypertrophic chondrocytes in vitro.
RT J. Bone Miner. Res. 10:157-163(1995).
RN [2]
RP SEQUENCE OF 49-97.
RX MEDLINE-82007831; PubMed-6792200;
RA Carr S.A., Hauschka P.V., Blemann K.;
RT Gas chromatographic mass spectrometric sequence determination of
RT osteocalcin, a gamma-carboxyglutamic acid-containing protein from
RT chicken bone.
RL J. Biol. Chem. 256:9944-9950(1981).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- P-TM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U10578; AAA78809.1; -.
DR PIR: A03305; GECH.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; Vitr_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Bone; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 48
FT CHAIN 49 97 OSTEOCALCIN.

```

FT MOD\_RES 65 65 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 72 72 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 71 77 BY SIMILARITY.  
 SQ SEQUENCE 97 AA: 10707 MW: 76868585C327062 CRC64;

Query Match 65.7%; Score 186.5; DB 1; Length 97;  
 Best Local Similarity 81.4%; Pred. No. 1.2e-16;  
 Matches 35; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

OY 7 GAPVPPDLPREVCELPDDELADHIGFQEAIRRFYGPV 49  
 DB 58 GAP---PNPLEAREVCELPDDELADHIGFQEAIRRFYGPV 97

RESULT 11  
 OSTC\_XENLA STANDARD; PRT; 49 AA.  
 AC P40147;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96139691; PubMed=8567186;  
 RA Canella M., Williamson M.K., Price P.A.;  
 RT "Amino-acid sequence of bone Gla protein from the African clawed toad  
 RT Xenopus laevis and the fish Sparus aurata";  
 RL Int. J. Pept. Protein Res. 46:419-423(1995).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTV: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 CC InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR002384; VltK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA; 1.  
 DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.  
 DR Calcium-binding; Gamma-carboxyglutamic acid; vitamin K; Bone.  
 FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 23 29 BY SIMILARITY.  
 SQ SEQUENCE 49 AA: 5360 MW: 7A9A6F3A12B6047 CRC64;

Query Match 65.5%; Score 186; DB 1; Length 49;  
 Best Local Similarity 94.3%; Pred. No. 6.2e-17;  
 Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 15 PLEPREVCELPDDELADHIGFQEAIRRFYGPV 49  
 DB 15 PLESOREVCELPDDELADHIGFQEAIRRFYGPV 49

RESULT 12  
 OSTC\_MOUSE STANDARD; PRT; 95 AA.  
 AC P04641;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)

DE (Bone Gla-protein) (BGP).  
 GN BGLAP1 AND BGLAP2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87004555; PubMed=3019668;  
 RA Celeste A.J., Buecker J.L., Kriz R., Wang E.A., Mooney J.M.;  
 RT "Isolation of the human gene for bone gla protein utilizing mouse and  
 RT rat cDNA clones";  
 RL EMBO J. 5:1885-1890(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94056292; PubMed=8243336;  
 RA Rahman S., Oberdorf A., Montecino M., Tanhauser S.M., Lian J.B.,  
 RA Stein G.S., Laipis P.J., Stein J.L.;  
 RT "Multiple copies of the bone-specific osteocalcin gene in mouse and  
 RT rat";  
 RL Endocrinology 133:3050-3053(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94117426; PubMed=8288580;  
 RA Desbois C., Hogue D.A., Karsenty G.;  
 RT "The mouse osteocalcin gene cluster contains three genes with two  
 RT separate spatial and temporal patterns of expression";  
 RL J. Biol. Chem. 269:1183-1190(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Calvaria;  
 RA Yotov W.V., St Arnaud R.;  
 RL Submitted (Jul-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- TISSUE SPECIFICITY: BONE.  
 CC -1- PTV: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES CODING FOR OSTEOCALCIN, THEIR  
 CC CODING SEQUENCE ONLY DIFFERS IN THE SIGNAL PEPTIDE REGION.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 DR EMBL: X01442; CAA27762.1; -  
 DR EMBL: S67455; AAB29145.1; -  
 DR EMBL: L24429; AAA39854.1; -  
 DR EMBL: L24431; AAA39856.1; -  
 DR EMBL: U11542; AAB40035.1; -  
 DR PIR: B25471; B25471.  
 DR MGD: MGI:88156; Bglapl.  
 DR MGD: MGI:88157; Bglap2.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR002384; VltK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA; 1.  
 DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.  
 DR Calcium-binding; Gamma-carboxyglutamic acid; vitamin K; Bone; Signal;  
 KW Multigene family.  
 FT SIGNAL 1 23 PROBABLE.  
 FT PROPEP 24 49 PROBABLE.  
 FT CHAIN 50 95 OSTEOCALCIN.  
 FT MOD\_RES 62 62 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).

```

FT  MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID
FT  MOD_RES 69 69 (BY SIMILARITY)
FT  MOD_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID
FT  DISULFID 68 74 (BY SIMILARITY)
FT  VARIANT 4 5 LS -> IF (IN BGLAP2)
FT  VARIANT 11 11 A -> T (IN BGLAP2)
FT  CONFLICT 23 23 A -> P (IN REF. 1)
SQ  SEQUENCE 95 AA: 10459 MW: 244611134805098 CRC64:
Query Match 58.5%; Score 166; DB 1; Length 95;
Best Local Similarity 65.2%; Pred. No. 4.2e-14;
Matches 30; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 LYOMLGAPVPPPLEPRREVCENLPCDELADHIGFOEAYRFFYG 47
Db 47 LRRYLGA SVSPDLEPTRELCENLPACDELSDQYGLKTRVRRYIG 92

RESULT 13
OSTR.MOUSE STANDARD: PRT; 95 AA.
AC P54615;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin-related protein precursor (OC-X) (Nephrocalcin).
GN BGLAP-RSL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94117426; PubMed=8288580;
RA Desbois C., Hogue D.A., Karsenty G.;
RT "The mouse osteocalcin gene cluster contains three genes with two
RT separate spatial and temporal patterns of expression.";
RT J. Biol. Chem. 269:1183-1190(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94062692; PubMed=8243336;
RA Rahman S., Oberdorf A., Montecino M., Tanhauser S.M., Llan J.B.,
RA Stein G.S., Lalpis P.J., Stein J.L.;
RT "Multiple copies of the bone-specific osteocalcin gene in mouse and
RT rat.";
RT Endocrinology 133:3050-3053(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ; TISSUE=Liver;
RA Yotou W.V., St. Arnaud R.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY BUT NOT IN BONE.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L24430; AAA39855.1; -
CC EMBL: S67456; AAB29146.1; -
CC EMBL: U11541; AAB60445.1; -
CC MGD: MGI:88155; Bglap-rsl.

```

```

DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla_1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU CARBOXYLATION; 1.
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Bone; Signal;
KW Multigene family.
FT SIGNAL 1 23 PROBABLE.
FT PROPEP 24 49 OSTEOCALCIN-RELATED PROTEIN.
FT CHAIN 50 95 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 62 62 (BY SIMILARITY).
FT MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 69 69 (BY SIMILARITY).
FT DISULFID 68 74 GAMMA-CARBOXYGLUTAMIC ACID
SQ SEQUENCE 95 AA: 10459 MW: 231968AA8B5848E CRC64:
Query Match 57.0%; Score 162; DB 1; Length 95;
Best Local Similarity 63.0%; Pred. No. 1.3e-13;
Matches 29; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 2 LYOMLGAPVPPPLEPRREVCENLPCDELADHIGFOEAYRFFYG 47
Db 47 LRRYLGA SVSPDLEPTRELCENLPACDELSDQYGLKTRVRRYIG 92

RESULT 14
OSTC.XIPGL STANDARD: PRT; 47 AA.
AC P02823;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
OS Xiphias gladius (swordfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Scombroidei;
OX Xiphidae; Xiphias.
OX NCBI_TaxID=8245;
RN [1]
RP SEQUENCE.
RA Price P.A., Otsuka A.S., Poser J.W.;
RT "Comparison of gamma-carboxyglutamic acid-containing proteins from
RT bovine and swordfish bone: primary structure and Ca++ binding.";
RL (In) Wasserman R.H., Corradino R.A., Caratoll E., Kretzinger R.H.,
RL MacLennan D.H., Siegel F.L. (eds.);
RL Calcium-binding proteins and calcium function, pp.333-337,
RL Elsevier, New York (1977).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC -----
CC PIR: A03306; GEMF.
CC InterPro: IPR002384; GLA_bone.
CC InterPro: IPR000294; VitK_dep_GLA.
CC Pfam: PF00594; gla_1.
CC PRINTS: PR00002; GLABONE.
CC PROSITE: PS00011; GLU CARBOXYLATION; 1.
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Bone.
FT MOD_RES 13 13 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 20 20 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 19 25
SQ SEQUENCE 47 AA: 5080 MW: 8093FE1787B09223 CRC64:

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:37:45 ; Search time 19 Seconds  
(Without alignments)  
247.926 Million cell updates/sec

Title: US-09-462-931-2

Sequence: 1 YLYQWLGAVPYPDPLEPRR.....DELADHIGFQEARFRYGPV 49

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR-73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	100	1 GEHU	osteocalcin precur
2	30	61.2	48	2 SOZ208	osteocalcin - emu
3	30	61.2	49	1 GEMKI	osteocalcin - crab
4	30	61.2	100	1 GEBQ	osteocalcin precur
5	17	34.7	49	1 GECT	osteocalcin - cat
6	15	30.6	49	1 A61280	osteocalcin - rabb
7	14	28.6	97	1 GECH	osteocalcin precur
8	11	22.4	99	1 GERT	osteocalcin precur
9	8	16.3	536	2 T42606	osteocalcin precur
10	7	14.3	47	1 GEMF	probable transcript
11	7	14.3	164	2 PC4154	osteocalcin - swor
12	7	14.3	275	2 E97419	osteocalcin - 2-glycop
13	7	14.3	286	2 B84226	hypothetical prote
14	7	14.3	317	2 G83544	3-hydroxyacyl-CoA
15	7	14.3	447	2 T38375	probable transcript
16	7	14.3	586	1 WMBETE	nuclear orphan rec
17	7	14.3	587	1 WMBETE	65k early nonstruc
18	7	14.3	660	2 AH2348	hypothetical prote
19	7	14.3	4767	2 T31345	hypothetical prote
20	6	12.2	70	2 C71823	ribosomal protein
21	6	12.2	70	2 B64590	ribosomal protein
22	6	12.2	95	2 B25471	osteocalcin precur
23	6	12.2	95	2 I53375	osteocalcin - mus
24	6	12.2	95	2 I67413	osteocalcin - relat
25	6	12.2	95	2 S61188	L71-1 protein - fr
26	6	12.2	100	2 S62333	yidd family lampor
27	6	12.2	103	2 E86564	yidd family - Chla
28	6	12.2	103	2 H72060	conserved hypotet
29	6	12.2	106	2 E81606	

30	6	12.2	119	2 S38261	myohemerithrin - p
31	6	12.2	132	2 T21416	hypothetical prote
32	6	12.2	136	2 T18052	DEAH box protein a
33	6	12.2	139	1 UDCB	cystatin precursor
34	6	12.2	143	2 T12144	hypothetical prote
35	6	12.2	169	2 S03744	hypothetical prote
36	6	12.2	171	2 S44737	C02C2.6 protein -
37	6	12.2	185	2 AH1075	probable fibrillar
38	6	12.2	210	2 S39645	acetoin utilizatio
39	6	12.2	210	2 C84054	acetoin dehydrogen
40	6	12.2	214	2 AB2392	hypothetical prote
41	6	12.2	235	2 D81201	cell division prot
42	6	12.2	242	2 D81776	cell division prot
43	6	12.2	252	2 AH3267	acetyltransferase
44	6	12.2	264	2 T35168	probable transcript
45	6	12.2	289	2 A87646	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

osteocalcin precursor [validated] - human

N:Alternate names: BGP; bone gla protein; gamma-carboxyglutamic acid-containing prote

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1980 #sequence, revision 07-Apr-1994 #text-change 08-Dec-2000

C:Accession: S12652; C25471; A03301; S08694

R:Kiefer, M.C.; Saphire, A.C.S.; Bauer, D.M.; Barr, P.J.

Nucleic Acids Res. 18, 1909, 1990

A:Title: The cDNA and derived amino acid sequences of human and bovine bone gla prote

A:Reference number: S12652; MUID:90245603; PMID:2336375

A:Accession: S12652

A:Molecule type: mRNA

A:Residues: 1-100 <KIE>

A:Cross-references: EMBL:X53698; NID:936092; PIDN:CAA37736.1; PID:936093

C:Caletke, A.J.; Rosen, V.; Buecker, J.L.; Kitz, R.; Wang, E.A.; Wozney, J.M.

EMBO J. 5, 1885-1890, 1986

A:Title: Isolation and sequence of the vitamin K-dependent protein from human bone. U

A:Reference number: A03301; MUID:81006914; PMID:6967872

A:Accession: A03301

A:Molecule type: protein

A:Residues: 52-100 <POS>

R:Calins, J.R.; Williamson, M.K.; Price, P.A.

Anal. Biochem. 199, 93-97, 1991

A:Title: Direct identification of gamma-carboxyglutamic acid in the sequencing of vit

A:Reference number: A44566; MUID:92222128; PMID:1807167

A:Contents: annotation

C:Comment: This protein, isolated from bone, binds strongly to apatite.

C:Comment: Alternative splicing may produce the sequence presented in reference A9104

C:Comment: Glu-68 is gamma-carboxylated in 9-50% of the molecules.

C:Genetics:

A:Gene: GDB:BGLAP

A:Cross-references: GDB:118760; OMIM:112260

A:Map position: 1q25-1q31

A:Introns: 22/1; 35/1; 58/2

C:Superfamily: osteocalcin

C:Keywords: bone; calcium binding; carboxyglutamic acid; extracellular matrix

F:1-51/Domain: signal sequence #status predicted <SIG>

F:52-100/Product: osteocalcin #status experimental <MAT>

F:60/Modified site: 4-Hydroxyproline (Pro) #status absent

F:68/Modified site: gamma-carboxyglutamic acid (Glu) (partial) #status experimental

F:72/75/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental

F:74-80/Disulfide bonds: #status experimental

Query Match 100.0%; Score 49; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-46;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWIGAVPPDPLEPRREVCENLPCDELADHIGFOEAYRRYGPV 49  
 |||  
 DB 52 YLYQWIGAVPPDPLEPRREVCENLPCDELADHIGFOEAYRRYGPV 100

RESULT 2  
 S02208  
 osteocalcin - emu  
 C:Species: Dromaius novaehollandiae (emu)  
 C>Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 18-Jun-1993  
 C:Accession: S02208

R:Huq, N.L.; Tseng, A.; Chapman, G.E.  
 Biochem. Int. 15, 271-277, 1987

A:Title: The amino acid sequence of emu osteocalcin: gas phase sequencing of gla-contain  
 A:Reference number: S02208; MUID:88134266; PMID:3501719  
 A:Accession: S02208  
 A:Molecule type: protein  
 A:Residues: 1-48 <HUQ>  
 C:Superfamily: osteocalcin

Query Match 61.2%; Score 30; DB 2; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-26;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 REVCELNPDCDELADHIGFOEAYRRYGPV 49  
 |||  
 DB 19 REVCELNPDCDELADHIGFOEAYRRYGPV 48

## RESULT 3

GEMKI  
 osteocalcin - crab-eating macaque

N:Alternate names: BGP; Bone Gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Macaca fascicularis (crab-eating macaque)  
 C>Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 06-Sep-1996

C:Accession: A03302  
 R:Hauschka, P.V.; Carr, S.A.; Blemann, K.  
 Biochemistry 21, 638-642, 1982

A:Title: Primary structure of monkey osteocalcin.  
 A:Reference number: A03302; MUID:82182842; PMID:6978733  
 A:Accession: A03302

A:Molecule type: protein  
 A:Residues: 1-49 <HAU>

C:Comment: This protein, isolated from bone, binds strongly to apatite.  
 C:Superfamily: osteocalcin  
 C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline

F:9/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:17,21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:23-29/Disulfide bonds: #status experimental

Query Match 61.2%; Score 30; DB 1; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-26;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 REVCELNPDCDELADHIGFOEAYRRYGPV 49  
 |||  
 DB 20 REVCELNPDCDELADHIGFOEAYRRYGPV 49

## RESULT 4

GRBO  
 osteocalcin precursor - bovine

N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 24-Apr-1984 #sequence\_revision 07-Apr-1994 #text\_change 22-Jun-1999

C:Accession: S12653; A03303; S08693  
 R:Kiefer, M.C.; Saphire, A.C.S.; Bauer, D.M.; Barr, P.J.  
 Nucleic Acids Res. 18, 1909, 1990

A:Title: The cDNA and derived amino acid sequences of human and bovine bone Gla protein

A:Reference number: S12652; MUID:90245603; PMID:2336375  
 A:Accession: S12653  
 A:Molecule type: mRNA

A:Residues: 1-100 <KIE>

A:Cross-References: EMBL:X53699; NID:q719; PIDN:CAA37737.1; PID:q720  
 A:Note: alternative splicing may produce a sequence lacking residues 33-34

R:Price, P.A.; Poser, J.W.; Raman, N.  
 Proc. Natl. Acad. Sci. U.S.A. 73, 3374-3375, 1976

A:Title: Primary structure of the gamma-carboxyglutamic acid-containing protein from  
 A:Reference number: A03303; MUID:77036749; PMID:1068450  
 A:Accession: A03303

A:Molecule type: protein  
 A:Residues: 52-100 <PRI>

C:Comment: This protein, isolated from bone, binds strongly to apatite.

C:Superfamily: osteocalcin

C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
 F:1-51/Domain: signal sequence #status predicted <SIG>

F:52-100/Product: osteocalcin #status experimental <MAT>

F:60/Modified site: 4-hydroxyproline (Pro) #status experimental

F:68,72,75/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:74-80/Disulfide bonds: #status experimental

Query Match 61.2%; Score 30; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-25;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 REVCELNPDCDELADHIGFOEAYRRYGPV 49  
 |||  
 DB 71 REVCELNPDCDELADHIGFOEAYRRYGPV 100

## RESULT 5

GECR

osteocalcin - cat

N:Alternate names: BGP; Bone Gla protein; gamma-carboxyglutamic acid-containing prote  
 C:Species: Felis silvestris catus (domestic cat)  
 C>Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 06-Sep-1996

C:Accession: A03304  
 R:Shimomura, H.; Kanai, Y.; Sanada, K.  
 J. Biochem. 96, 405-411, 1984

A:Title: Primary structure of cat osteocalcin.  
 A:Reference number: A03304; MUID:85054706; PMID:6334077  
 A:Accession: A03304

A:Molecule type: protein  
 A:Residues: 1-49 <SHI>

C:Superfamily: osteocalcin

C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
 F:9/Modified site: 4-hydroxyproline (Pro) #status experimental

F:17,21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:23-29/Disulfide bonds: #status experimental

Query Match 34.7%; Score 17; DB 1; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-11;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 CELNPDCDELADHIGFO 39  
 |||  
 DB 23 CELNPDCDELADHIGFO 39

## RESULT 6

A61280

osteocalcin - rabbit

N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing prote  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 12-May-1994 #sequence\_revision 02-Jun-1994 #text\_change 06-Sep-1996

C:Accession: A61280  
 R:Viridi, A.S.; Willis, A.C.; Hauschka, P.V.; Triffitt, J.T.  
 Biochem. Soc. Trans. 19, 373S, 1991

A:Title: Primary aminoacid sequence of rabbit osteocalcin.

A:Reference number: A61280; MUID:92175242; PMID:1794506  
 A:Accession: A61280

A:Molecule type: protein

A:Residues: 1-49 <VIR>  
 C:Superfamily: osteocalcin  
 C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline; pyroglutamic acid  
 F:1/Modified site: pyrrolidone carboxylic acid (Glu) #status experimental  
 F:9/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:17,21,24/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:23-29/Disulfide bonds: #status predicted

Query Match 30.6%; Score 15; DB 1; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1; 7e-09;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 REVCENPDCDELAD 34  
 Db 20 REVCENPDCDELAD 34

# RESULT 7 GECH

osteocalcin precursor - chicken  
 N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Gallus gallus (chicken)  
 C:Date: 15-Oct-1982 #sequence\_revision 15-Aug-1997 #text\_change 22-Jun-1999  
 C:Accession: I50700; A03305  
 R:Neugebauer, B.M.; Moore, M.A.; Broess, M.; Gerstenfeld, L.C.; Hauschka, P.V.  
 J. Bone Miner. Res. 10, 157-163, 1995  
 A:Title: Characterization of structural sequences in the chicken osteocalcin gene: expt  
 A:Reference number: I50700; MUID:95266465; PMID:77477623  
 A:Accession: I50700  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-97 <NEU>  
 A:Cross-references: EMBL:U10578; NID:g1008455; PIDN:AAA7809.1; PID:g595408  
 R:Carri, S.A.; Hauschka, P.V.; Biemann, K.  
 J. Biol. Chem. 256, 9944-9950, 1981  
 A:Title: Gas chromatographic mass spectrometric sequence determination of osteocalcin, a  
 A:Reference number: A03305; MUID:82007831; PMID:6792200  
 A:Accession: A03305  
 A:Molecule type: protein  
 A:Residues: 49-63, 'I', 65-77, 'N', 79-82, 'E', 85-90, 'Q', 91-97 <CAR>  
 C:Comment: The gamma-carboxyglutamic acid residues formed by vitamin K-dependent posttr  
 C:Superfamily: osteocalcin  
 C:Keywords: bone; calcium binding; carboxyglutamic acid  
 F:49-97/Product: osteocalcin #status experimental <MAT>  
 F:65,69,72/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
 F:71-77/Disulfide bonds: #status predicted

Query Match 28.6%; Score 14; DB 1; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 4e-08;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 36 IGFOEAVRRYGPV 49  
 Db 84 IGFOEAVRRYGPV 97

# RESULT 8 GERT

osteocalcin precursor - rat  
 N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing protein  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 22-Jun-1999  
 C:Accession: A31856; A31419; A32324; A25167; A25471  
 R:Toom, K.; Rutledge, S.J.C.; Buenaaga, R.F.; Rodan, G.A.  
 Biochemistry 27, 8521-8526, 1988  
 A:Title: Characterization of the rat osteocalcin gene: stimulation of promoter activity  
 A:Reference number: A31856; MUID:89118266; PMID:3265336  
 A:Accession: A31856  
 A:Molecule type: DNA  
 A:Residues: 1-99 <TOO>  
 A:Cross-references: GB:M23637; NID:g340986; PIDN:AAA1761.1; PID:g514962  
 R:Theofan, G.; Haberstroh, L.M.; Price, P.A.  
 DNA 8, 213-221, 1989

A:Title: Molecular structure of the rat bone Gla protein gene and identification of p  
 A:Reference number: A31419; MUID:89251082; PMID:2785907

A:Accession: A31419  
 A:Molecule type: DNA  
 A:Residues: 1-99 <THE>  
 A:Cross-references: GB:M25490; NID:g576530; PIDN:AAA3280.1; PID:g576531  
 R:Ilari, J.; Stewart, C.; Puchacz, E.; Mackowiak, S.; Shalhoub, V.; Collart, D.; Zame  
 Proc. Natl. Acad. Sci. U.S.A. 86, 1143-1147, 1989  
 A:Title: Structure of the rat osteocalcin gene and regulation of vitamin D-dependent  
 A:Reference number: A32324; MUID:89145200; PMID:2784002  
 A:Accession: A32324  
 A:Molecule type: DNA  
 A:Residues: 1-99 <LID>  
 A:Cross-references: GB:J04500; NID:g205863; PIDN:AAA1764.1; PID:g205864  
 R:Pan, L.C.; Price, P.A.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 6109-6113, 1985

A:Title: The propeptide of rat bone gamma-carboxyglutamic acid protein shares homolog  
 A:Reference number: A25167; MUID:85298305; PMID:3875856  
 A:Accession: A25167  
 A:Molecule type: mRNA  
 A:Residues: 1-99 <PAN>  
 A:Cross-references: GB:M11777; NID:g203147; PIDN:AAA0816.1; PID:g203148  
 R:Ceslote, A.J.; Rosen, V.; Buecker, J.L.; Kriz, R.; Wang, E.A.; Mooney, J.M.  
 EMBO J. 5, 1885-1890, 1986

A:Title: Isolation of the human gene for bone gla protein utilizing mouse and rat cdn  
 A:Reference number: A91045; MUID:87004555; PMID:3019668  
 A:Accession: A25471  
 A:Molecule type: mRNA  
 A:Residues: 1-99 <CEL>  
 A:Cross-references: GB:X04141; NID:g55826; PIDN:CAA27761.1; PID:g55827  
 C:Genetics:  
 A:Introns: 22/1; 33/1; 56/2

C:Superfamily: osteocalcin  
 C:Keywords: bone; calcium binding; carboxyglutamic acid; hydroxyproline  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-49/Domain: propeptide #status predicted <PRO>  
 F:50-99/Product: osteocalcin #status predicted <OCN>  
 F:58/Modified site: 4-hydroxyproline (Pro) #status predicted  
 F:66,70,73/Modified site: gamma-carboxyglutamic acid (Glu) #status predicted  
 F:72-78/Disulfide bonds: #status predicted

Query Match 22.4%; Score 11; DB 1; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 8e-05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 29 CDELADHIGFQ 39  
 Db 78 CDELADHIGFQ 88

# RESULT 9 T42606

probable transcription activator - equine herpesvirus 4 (strain NS80567)  
 C:Species: equine herpesvirus 4  
 A:Variety: strain NS80567  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000  
 C:Accession: T42606  
 R:Telord, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.  
 J. Gen. Virol. 79, 1197-1203, 1998  
 A:Title: The DNA sequence of equine herpesvirus-4.  
 A:Reference number: 222173; MUID:98264497; PMID:9603335  
 A:Accession: T42606  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-536 <TEL>  
 A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59582.1; PID:g2606010  
 A:Experimental source: strain NS80567  
 C:Genetics:  
 A:Note: 63  
 C:Superfamily: varicella-zoster virus gene 63 protein; RING finger homology  
 F:5-53/Domain: RING finger homology <RRN>

Query Match 16.3%; Score 8; DB 2; Length 536;

Best Local Similarity 100.0%; Pred. No. 0.72;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 PVPDPPL 16  
| | | | | | |  
DB 486 PVPDPPL 493

# RESULT 10

GENE

osteocalcin - swordfish

N:Alternate names: BGP; bone Gla protein; gamma-carboxyglutamic acid-containing protein  
C:Species: *Xiphias gladius* (swordfish)  
C:Date: 30-Apr-1979 #sequence\_revision 27-Nov-1985 #text\_change 06-Sep-1996  
C:Accession: A03306

R:Price, P.A.; Otsuka, A.S.; Poser, J.W.  
In Calcium-binding Proteins and Calcium Function, Wasserman, R.H., Corradino, R.A., Carr  
A:title: Comparison of gamma-carboxyglutamic acid-containing proteins from bovine and sw  
A:Reference number: A03306  
A:Accession: A03306

A:Molecule type: protein  
A:Residues: 1-47 <PR>

A>Note: residues 14, 24, and 37 were not positively identified  
C:Superfamily: osteocalcin  
C:Keywords: bone; calcium binding; carboxyglutamic acid  
F:13,17,20/Modified site: gamma-carboxyglutamic acid (Glu) #status experimental  
F:19-25/Disulfide bonds: #status predicted

Query Match 14.3%; Score 7; DB 1; Length 47;  
Best Local Similarity 100.0%; Pred. No. 0.97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 REVCELN 26  
| | | | | | |  
DB 16 REVCELN 22

# RESULT 11

PC4154

Zn-alpha 2-glycoprotein - rat (fragment)

C:Species: *Rattus norvegicus* (Norway rat)

C:Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 11-Jan-2000  
C:Accession: PC4154

R:Urra, J.A.; Fueyo, A.; Balbin, M.; Velasco, G.; Pendas, A.M.; Lopez-Otin, C.  
Gene 169, 233-236, 1996

A:title: Alternative splicing gives rise to two novel long isoforms of Zn-alpha2-glycopo  
A:Reference number: PC4154; MUID:96194808; PMID:8647453

A:Accession: PC4154

A:Molecule type: mRNA

A:Residues: 1-164 <UR>

A:Cross-references: EMBL:X86178

A:Experimental source: liver

C:Comment: It is involved in the transport of nonpolymorphic substances or intercellular  
n.

C:Genetics:

A:Gene: Zn-alpha 2-gp

A:introns: 7/1; 77/2; 83/1

C:Superfamily: class I histocompatibility antigen: immunoglobulin homology

C:Keywords: glycoprotein; immune response; immunoglobulin; liver

F:79/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 14.3%; Score 7; DB 2; Length 164;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 PVPDPPL 17  
| | | | | | |  
DB 48 PVPDPPL 54

# RESULT 12

ES97419

hypothetical protein AGR\_C-880 [imported] - *Agrobacterium tumefaciens* (strain C58, Cered

C:Species: *Agrobacterium tumefaciens*

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002

C:Accession: E97419

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm  
A.; Liu, F.; Woliam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,  
Science 294, 2323-2328, 2001

A:title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium*

A:Reference number: A97359; PMID:11743194

A:Accession: E97419

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK86310.1; PID:q15155426; GSPDB:GN00169

C:Genetics:

A:Gene: AGR\_C-880

A:Map position: circular chromosome

Query Match 14.3%; Score 7; DB 2; Length 275;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 FOEAYRR 44  
| | | | | | |  
DB 167 FOEAYRR 173

# RESULT 13

B84226

3-hydroxyacyl-CoA dehydrogenase [imported] - *Halobacterium* sp. NRC-1

C:Species: *Halobacterium* sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: B84226

R:Ng, W.Y.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky  
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja  
Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;  
A:title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: B84226

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <STO>

A:Cross-references: GB:AE004437; NID:q10580268; PIDN:AA619174.1; GSPDB:GN00138

C:Genetics:

A:Gene: hbd1

C:Superfamily: 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenase homol

Query Match 14.3%; Score 7; DB 2; Length 286;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 ELADHTG 37  
| | | | | | |  
DB 234 ELADHTG 240

# RESULT 14

G83544

probable transcription regulator PA0815 [imported] - *Pseudomonas aeruginosa* (strain P

C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: G83544

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: G83544

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-317 <STO>

A:Cross-references: GB:AE004516; GB:AE004091; NID:g9946699; PIDN:AAG04204.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA0815

Query Match 14.3%; Score 7; DB 2; Length 317;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 OMLGAPV 10  
|||||  
DB 42 OMLGAPV 48

## RESULT 15

138975  
nuclear orphan receptor LXR-alpha - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Feb-1996 #sequence\_revision 16-Feb-1996 #text\_change 20-Sep-1999  
C:Accession: 138975  
R:Willy, P.J.; Umesono, K.; Ong, E.S.; Evans, R.M.; Heyman, R.A.; Mangelsdorf, D.J.  
Genes Dev. 9, 1033-1045, 1995  
A:Title: LXR, a nuclear receptor that defines a distinct retinoid response pathway.  
A:Reference number: 138975; MUID:95262897; PMID:7744246  
A:Accession: 138975  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-447 <RES>  
A:Cross-references: EMBL:U22662; NID:g726512; PIDN:AAA85856.1; PID:g726513  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: zinc finger  
F:96-367/Domain: erba transforming protein homology <ERRA>

Query Match 14.3%; Score 7; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 WIGAPVP 11  
|||||  
DB 4 WIGAPVP 10

## RESULT 16

## KMBEDE

65K early nonstructural protein - human cytomegalovirus (strain AD169)  
N:Alternate names: UL84 protein  
C:Species: human cytomegalovirus, human herpesvirus 5  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 16-Jul-1999  
C:Accession: S09848  
R:Chen, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Hornsneil, T.;  
M.; Barrett, B.G.  
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
A:Reference number: S09749; MUID:90269039; PMID:2161319  
A:Accession: S09848  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-586 <CHE>  
A:Cross-references: EMBL:X17403; NID:g959591; PIDN:CAA5358.1; PID:g1780862  
A:Note: this sequence was submitted to the EMBL Data Library, December 1989  
C:Superfamily: human cytomegalovirus 65K early nonstructural protein  
C:Keywords: leucine zipper; nonstructural protein  
F:114-135/Region: leucine zipper motif  
F:324-372/Region: leucine zipper motif

Query Match 14.3%; Score 7; DB 1; Length 586;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYGP 48  
|||||  
DB 443 YRRFYGP 449

## RESULT 17

## KMBETE

65K early nonstructural protein - human cytomegalovirus (strain Towne)  
C:Species: human cytomegalovirus, human herpesvirus 5  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 16-Jul-1999  
C:Accession: A41808  
R:He, Y.S.; Xu, L.; Huang, E.S.  
J. Virol. 66, 1098-1108, 1992  
A:Title: Characterization of human cytomegalovirus UL84 early gene and identification  
A:Reference number: A41808; MUID:92114132; PMID:1309892  
A:Accession: A41808  
A:Molecule type: mRNA  
A:Residues: 1-587 <HE>  
A:Cross-references: GB:M81432; NID:g330541; PIDN:AAA45947.1; PID:g330542  
C:Superfamily: human cytomegalovirus 65K early nonstructural protein  
C:Keywords: leucine zipper; nonstructural protein  
F:114-135/Region: leucine zipper motif  
F:325-373/Region: leucine zipper motif

Query Match 14.3%; Score 7; DB 1; Length 587;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRFYGP 48  
|||||  
DB 444 YRRFYGP 450

## RESULT 18

## AH2348

hypothetical protein al14343 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp.  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C:Accession: AH2348  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Itigun  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata  
DNA Res. 8, 205-213, 2001  
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH2348  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-660 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA876042.1; PID:g17133479; GSPDB:GN00179  
C:Genetics:  
A:Gene: al14343

Query Match 14.3%; Score 7; DB 2; Length 660;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 PDPLEPR 19  
|||||  
DB 312 PDPLEPR 318

## RESULT 19

## T31345

hypothetical protein G01D9.5 - Caenorhabditis briggsae  
C:Species: Caenorhabditis briggsae  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T31345  
R:Waterston, R.  
submitted to the EMBL Data Library, April 1996  
A:Description: The C. briggsae genome sequencing project.  
A:Reference number: Z21010  
A:Accession: T31345  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-4767 <MAT>  
A:Cross-references: EMBL:U56248; NID:g1293789; PID:g1293790; PIDN:AAA98699.1  
C:Genetics:  
A:Introns: 269/3; 341/3; 853/1; 920/2; 4452/3; 4534/3; 4592/3; 4654/3; 4670/3; 4707/2; 4  
A>Note: G01D9.5  
C:Superfamily: acyl carrier protein homology; acetate-CoA ligase homology  
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein  
F:3472-3541/Domain: acyl carrier protein homology <ACPL>  
F:4039-4427/Domain: acetate-CoA ligase homology <ACPL>  
F:4447-4514/Domain: acyl carrier protein homology <ACP2>  
F:2210.3505/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 14.3%; Score 7; DB 2; Length 4767;  
Best Local Similarity 100.0%; Pred. No. 65;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 DELADHI 36  
|111111|  
Db 35 DELADHI 41

## RESULT 20

ribosomal protein S21 - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 24-Sep-1999

C:Accession: C71923

A:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: C71923

A:Molecule type: DNA

A:Residues: 1-70 <ARN>

A:Cross-references: GB:AE001484; GB:AE001439; NID:94155043; PIDN:AAD06085.1; PID:9415504

A:Experimental source: strain J99

C:Genetics:

A:Gene: rpsu  
C:Superfamily: Escherichia coli ribosomal protein S21

Query Match 12.2%; Score 6; DB 2; Length 70;  
Best Local Similarity 100.0%; Pred. No. 17;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EAYRRF 45  
|11111|  
Db 14 EAYRRF 19

## RESULT 21

ribosomal protein S21 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 13-Aug-1999

C:Accession: B64590

R:Tom, J.F.; White, O.; Keilavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKen

son, J.D.; Kelley, J.M.; Cotton, M.D.; Meldman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.  
Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: B64590

A:Molecule type: DNA  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Residues: 1-70 <TOM>  
A:Cross-references: GB:AE005070; GB:AE005011; NID:92313672; PIDN:AAD07628.1; PID:9231367  
C:Superfamily: Escherichia coli ribosomal protein S21

Query Match 12.2%; Score 6; DB 2; Length 70;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EAYRRF 45  
|11111|  
Db 14 EAYRRF 19

## RESULT 22

osteocalcin precursor - mouse  
C:Species: Mus musculus (house mouse)

C:Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 20-Aug-1999

C:Accession: B25471; A49871; I61189

R:Celeste, A.J.; Rosen, V.; Buecker, J.L.; Kriz, R.; Wang, E.A.; Wozney, J.M.  
EMBO J. 5, 1885-1890, 1986

A:Title: Isolation of the human gene for bone gla protein utilizing mouse and rat cDN

A:Reference number: A91045; MUID:87004535; PMID:3019668

A:Accession: B25471

A:Molecule type: DNA

A:Residues: 1-95 <CEL>

A:Cross-references: GB:I24429; NID:9455452; PIDN:AAA39854.1; PID:9455453

A:Accession: I61189

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3,'1F',6-10,'T',12-95 <RES>

A:Cross-references: GB:I24431; NID:9455456; PIDN:AAA39856.1; PID:9455457

C:Genetics:

A:Introns: 22/1; 33/1; 52/2

C:Superfamily: osteocalcin

F:1-49/Domain: signal sequence #status predicted <SIG>  
F:50-95/Product: osteocalcin #status predicted <MAT>

Query Match 12.2%; Score 6; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 23;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 PDPLEP 18  
|11111|  
Db 58 PDPLEP 63

## RESULT 23

osteocalcin - mouse  
C:Species: Mus sp. (mouse)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 20-Aug-1999

C:Accession: I53275

R:Rahman, S.; Oberdorf, A.; Montecino, M.; Tanhauser, S.M.; Llan, J.B.; Stein, G.S.;  
Endocrinology 133, 3050-3053, 1993

A:Title: Multiple copies of the bone-specific osteocalcin gene in mouse and rat.

A:Reference number: I53275; MUID:94062692; PMID:8243336

A:Accession: I53275

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>  
A:Cross-references: GB:S67455; NID:9456854; PIDN:AA929145.1; PID:9456856

C:Genetics:  
A:Introns: 22/1; 33/1; 52/2; 72/2  
C:Superfamily: osteocalcin

Query Match 12.2%; Score 6; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 23;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	13	PDPLEP	18
Db	58	PDPLEP	63

RESULT 24  
167413  
osteocalcin - mouse  
C:Species: Mus sp. (mouse)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Aug-1999  
C:Accession: 167413  
R:Rauman, S.; Oberdorf, A.; Montecino, M.; Tanhauser, S.M.; Iian, J.B.; Stein, G.S.; Laitinen, J.  
Endocrinology 133, 3050-3053, 1993  
A:Title: Multiple copies of the bone-specific osteocalcin gene in mouse and rat.  
A:Reference number: 153275; MUID:94062692; PMID:824336  
A:Accession: 167413  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-95 <RES>  
A:Cross-references: GB:S67456; NID:g456855; PIDN:AMB29146.1; PID:g456857  
C:Genetics:  
A:Introns: 22/1, 33/1, 52/2, 72/2  
C:Superfamily: osteocalcin

```

RESULT 25
161188
osteocalcin-related protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C:Accession: 161188; 149073
R:Desbois, C.; Hoque, P. A.; Karsenty, G.
J. Biol. Chem. 269, 1183-1190, 1994
A:Title: The mouse osteocalcin gene cluster contains three genes with two separate start
A:Reference number: A49871; MUID:94117426; PMID:8288580
A:Accession: 161188
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-95 <RES>
A:Cross-references: GB:I24430; NID:g455454; PIDN:AAA38655.1; PID:g455455
A:Accession: 149073
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-95 <RES>
A:Cross-references: EMBL:U11541; NID:g508297; PIDN:AA860445.1; PID:g508298
C:Genetics:
A:Introns: 22/1; 33/1; 52/2
C:Superfamily: osteocalcin

Query Match          12.2%  Score 6;  DB 2;  Length 95;
Best Local Similarity 100.0%;  Pred. No. 23;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      13 PDPLEP 18
      |||||
Db      58 PDPLEP 63

RESULT 26
S62333
L71-1 protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999
C:Accession: S62333;

```

R:Wright, L.G.,Chen, T.; Thummel, C.S.; Guild, G.M.  
J. Mol. Biol. 255, 387-400, 1996  
A:Title: Molecular characterization of the 71E late puff in *Drosophila melanogaster*  
A:Reference number: S62333; MUID:56152797; PMID:8568884  
A:Accession: S62333  
A:Molecule type: DNA  
A:Residues: 1-100 <MEM>  
A:Cross-references: EMBL:U023836; NID:g939996; PIDN:AAA74176.1; PID:g939997  
A:Accession: S62342  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-100 <MEM>  
A:Cross-references: EMBL:U04095; NID:g775224; PIDN:AAA65109.1; PID:g775225  
C:Genetics:  
A:Gene: L71-1  
A:Cross-references: Flybase:FBgn0004588  
A:Introns: 12/1; 79/1  
C:Superfamily: L71-10 protein

```

RESULT 27
E86564
Yidd family [imported] - Chlamydophila pneumoniae (strain J138)
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: E86564
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: E86564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-103 <Sto>
A:Cross-references: GB:BA000008; NID:g8978964; PIDN:BA9879.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: yidd

Query Match 12.2%; Score 6; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWL 6
|||||
DB 20 YLYQWL 25

RESULT 28
H72060
Yidd family - Chlamydophila pneumoniae (strain CW1029)
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: H72060
R:Kalmán, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: H72060
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-103 <ARN>
A:Cross-references: GB:AE001643; GB:AE001365; NID:g4376876; PIDN:AAD18731.1; PID:g437
A:Experimental source: strain CW1029
C:Genetics:

```

A:Gene: yidd

Query Match 12.2%; Score 6; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWL 6  
|||||  
DB 20 YLYQWL 25

RESULT 29

E81606  
conserved hypothetical protein CP0156 [imported] - Chlamydia pneumoniae (strain AR39)  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: E81606

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: E81606

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-106 <REA>  
A:Cross-references: GB:AE002177; GB:AE002161; NID:g7189090; PIDN:AAF38037.1; PID:g718909  
A:Experimental source: strain AR39, HL cells  
C:Genetics:

A:Gene: CP0156

Query Match 12.2%; Score 6; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWL 6  
|||||  
DB 23 YLYQWL 28

RESULT 30

S38261  
myohemerythrin - polychaete (Nereis diversicolor)  
C:Species: Nereis diversicolor (sandworm)  
C>Date: 19-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 22-Oct-1999  
C:Accession: S38261

R:Demyanok, S.; Li, K.W.; van der Schors, R.; Dhainaut-Courtois, N.  
Eur. J. Biochem. 217, 151-156, 1993

A:Title: Amino acid sequence of the small cadmium-binding protein (MP II) from Nereis diversicolor.  
A:Reference number: S38261; MUID:94039032; PMID:823553  
A:Accession: S38261

A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-119 <DEM>  
C:Superfamily: hemerythrin  
C:Keywords: iron; oxygen carrier  
F:25,54,58,74,78,107,112/Binding site: 2Fe-O cluster (His, His, Glu, His, His, His, Asp)

Query Match 12.2%; Score 6; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11  
|||||  
DB 88 LGAPVP 93

RESULT 31

T21416  
hypothetical protein F26E4.9 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T21416

R:Lightning, J.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19419  
A:Accession: T21416  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-132 <MIL>  
A:Cross-references: EMBL:Z81070; PIDN:CAB03002.1; GSPDB:GN00019; CESP:F26E4.9  
A:Experimental source: clone F26E4  
C:Genetics:

A:Gene: CESP:F26E4.9  
A:Map position: 1  
A:Introns: 62/3; 96/1  
C:Superfamily: mammalian cytochrome-c oxidase chain Vb

Query Match 12.2%; Score 6; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 YDPPLE 17  
|||||  
DB 39 YDPPLE 44

RESULT 32

T18052  
DEAH box protein a550R - Chlorella virus PBCV-1  
C:Species: Chlorella virus PBCV-1  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T18052

R:Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z18806  
A:Accession: T18052

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-136 <GRA>  
A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96997.1  
A:Experimental source: specific host Chlorella strain NC64A  
C:Genetics:

A:Note: a550R  
C:Superfamily: Chlorella virus PBCV-1 DEAH box protein a550R

Query Match 12.2%; Score 6; DB 2; Length 136;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVP 11  
|||||  
DB 12 LGAPVP 17

RESULT 33

UDCH  
cystatin precursor - chicken  
N:Alternate names: cystatin 1; cysteine proteinase inhibitor; egg-white cystatin  
C:Species: Gallus gallus (chicken)  
C>Date: 03-Aug-1984 #sequence\_revision 12-Apr-1996 #text\_change 29-Oct-1999  
C:Accession: A34456; A01274; S01461; S48159; S04008; JN0789

R:Collella, R.; Sakeguchi, Y.; Nagase, H.; Bird, J.W.C.  
J. Biol. Chem. 264, 17164-17169, 1989

A:Title: Chicken egg white cystatin. Molecular cloning, nucleotide sequence, and tissue distribution.  
A:Reference number: A34456; MUID:90008873; PMID:2793849  
A:Accession: A34456

A:Molecule type: mRNA  
A:Residues: 1-139 <COI>  
A:Cross-references: GB:J05077; NID:g211714; PIDN:AAA48744.1; PID:g211715  
R:Schwabe, C.; Anastasi, A.; Crow, H.; McDonald, J.R.; Barrett, A.J.  
Biochem. J. 217, 813-817, 1984

A:Title: Cystatin. Amino acid sequence and possible secondary structure.  
A:Reference number: A01274; MUID:84178305; PMID:6712597  
A:Accession: A01274

A:Molecule type: protein



A:Residues: 24-139 <SCH>  
 R:Turk, V.; Brzin, J.; Longfer, M.; Ritonja, A.; Eropkin, M.; Borchardt, U.; Machleidt, W.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983  
 A:Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatins  
 A:Reference number: 501461; PMID:84110059; PMID:6662498  
 A:Accession: 501461  
 A:Molecule type: protein  
 A:Residues: 24-139 <TUR>  
 R:Anastasi, A.; Brown, M.A.; Kembhavi, A.A.; Nicklin, M.J.H.; Sayers, C.A.; Sunter, D.C.  
 Biochem. J. 211, 129-138, 1993  
 A:Title: Cystatin, a protein inhibitor of cysteine proteinases. Improved purification for  
 A:Reference number: A37514; PMID:83256421; PMID:6409085  
 A:Contents: annotation; characterization of protein  
 R:Grubb, A.; Lofberg, H.; Barrett, A.J.  
 FEBS Lett. 170, 370-374, 1984  
 A:Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.  
 A:Reference number: 501462  
 A:Contents: annotation; disulfide bonds  
 R:Auerswald, E.A.; Naegler, D.K.; Schulze, A.J.; Engh, R.A.; Genenger, G.; Machleidt, W.  
 Eur. J. Biochem. 224, 407-415, 1994  
 A:Title: Production, inhibitory activity, folding and conformational analysis of an N-ter  
 A:Reference number: 548159; PMID:95010016; PMID:7925354  
 A:Accession: 548159  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 24-139 <AUE>  
 R:Labber, B.; Kriegstein, K.; Henschen, A.; Kos, J.; Turk, V.; Huber, R.; Bode, W.  
 FEBS Lett. 248, 162-168, 1989  
 A:Title: The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein.  
 A:Reference number: 504008; PMID:89252033; PMID:2721673  
 A:Accession: 504008  
 A:Molecule type: protein  
 A:Residues: 97-114 <LAB>  
 R:Colletta, R.; Bird, J.W.C.  
 Gene 130, 175-181, 1993  
 A:Title: Isolation and characterization of the chicken cystatin-encoding gene: Mapping t  
 A:Reference number: JN0789; PMID:93366172; PMID:8359684  
 A:Accession: JN0789  
 A:Molecule type: DNA  
 A:Residues: 1-139 <CO2>  
 A:Cross-references: GB:M95725  
 A:Note: authors failed to translate the codon for residue 115-Tyr  
 C:Comment: This protein binds tightly to and inhibits a variety of cysteine proteinases  
 C:Genetics:  
 A:Gene: Csn  
 A:Introns: 76/3; 114/3  
 C:Superfamily: cystatin; cystatin homology  
 C:Keywords: cysteine proteinase inhibitor; egg white; phosphoprotein  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-139/Product: cystatin, long form #status experimental <CYLF>  
 F:30-139/Domain: cystatin homology <CYS>  
 F:32-139/Product: cystatin, short form #status experimental <CYSF>  
 F:76-80/Region: inhibitory #status predicted  
 F:94-104,118-138/Disulfide bonds: #status experimental  
 F:103/Binding site: phosphate (ser) (covalent) (partial) #status experimental  
 Query Match 12.2%; Score 6; DB 1; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 LGAPVP 11  
 DB 31 LGAPVP 36  
 RESULT 34  
 T12144  
 hypothetical protein 143 - fava bean mitochondrion  
 C:Species: mitochondrion Vicia faba (fava bean)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Sep-1999  
 C:Accession: T12144  
 R:Schepers, D.G.; Iino, H.; Boutry, M.  
 Plant Sci. 129, 203-212, 1997

A:Title: Variant mitochondrial transcripts of a broad bean line are associated with t  
 A:Reference number: 217435  
 A:Accession: T12144  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-143 <SCH>  
 A:Cross-references: EMBL:J36945; NID:9558142; PID:9558143  
 C:Genetics:  
 A:Genome: mitochondrion  
 C:Superfamily: fava bean mitochondrion hypothetical protein 143  
 C:Keywords: mitochondrion  
 Query Match 12.2%; Score 6; DB 2; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 APVPYP 13  
 DB 55 APVPYP 60  
 RESULT 35  
 S03744  
 hypothetical protein - Calothrix sp.  
 C:Species: Calothrix sp.  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 18-Jun-1993  
 C:Accession: S03744  
 R:Parsof, C.; Mazel, D.  
 Mol. Microbiol. 1, 45-52, 1987  
 A:Title: Cloning and nucleotide sequence of the thrp gene from the cyanobacterium Cal  
 A:Reference number: S03743; PMID:88260883; PMID:2838727  
 A:Accession: S03744  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-169 <PAR>  
 Query Match 12.2%; Score 6; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 40 EAYRRP 45  
 DB 43 EAYRRP 48  
 RESULT 36  
 S44737  
 C02C2.6 protein - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Sep-1997  
 C:Accession: S44737  
 R:Wilson, R.  
 submitted to the EMBL Data Library, September 1993  
 A:Description: Sequence of the C. elegans cosmid C02C2.  
 A:Reference number: S44737  
 A:Accession: S44737  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-171 <WIL>  
 A:Cross-references: EMBL:L23649; NID:9388566; PID:9388567  
 C:Genetics:  
 A:Introns: 79/3; 118/3  
 Query Match 12.2%; Score 6; DB 2; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 30 DELADH 35  
 DB 109 DELADH 114  
 RESULT 37



rl, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masigian, V.; Pizze, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V  
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A:Reference number: AB1000; MUID:2015755; PMID:10710307  
A:Accession: D81201  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1235 <TET>  
A:Cross-references: GB:AE002398; GB:AE002098; NID:97225640; PIDN:AAF40863.1; PID:9722564  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB0425  
C:Superfamily: cell division protein ftsQ

Query Match 12.2%; Score 6; DB 2; Length 235;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 QEAYR 44  
Db 70 QEAYR 75

RESULT 42  
cell division protein NMA2059 [imported] - *Neisseria meningitidis* (strain Z2491 serogrou  
C:Species: *Neisseria meningitidis*  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: G81776  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.  
A:Reference number: AB1775; MUID:2022556; PMID:10761919  
A:Accession: G81776  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1242 <PAR>  
A:Cross-references: GB:AL162758; GB:AL157959; NID:97380672; PIDN:CAB85277.1; PID:9738068  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: ftsQ; NMA2059  
C:Superfamily: cell division protein ftsQ

Query Match 12.2%; Score 6; DB 2; Length 242;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 QEAYR 44  
Db 77 QEAYR 82

RESULT 43  
acetyltransferase (EC 2.3.1.-) [imported] - *Brucella melitensis* (strain 16M)  
C:Species: *Brucella melitensis*  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AH3267  
R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitens*  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AH3267  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-252 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL51307.1; PID:917982002; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0125

A:Map position: I  
C:Keywords: acyltransferase

Query Match 12.2%; Score 6; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 37 GFOEAY 42  
Db 237 GFOEAY 242

RESULT 44  
T35168  
probable transcription regulator - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jan-2000  
C:Accession: T35168  
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, July 1998  
A:Reference number: Z21570  
A:Accession: T35168  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-264 <SRE>  
A:Cross-references: EMBL:AL031107; PIDN:CAA19948.1; GSPDB:GN00070; SCOEDB:SC5A7.19C  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC5A7.19C  
C:Superfamily: acetate operon repressor

Query Match 12.2%; Score 6; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 38 FOEAYR 43  
Db 153 FOEAYR 158

RESULT 45  
A87646  
hypothetical protein CC3203 [imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: A87646  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: A87646  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-289 <STO>  
A:Cross-references: GB:AE005673; NID:913424881; PIDN:AAK25165.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3203

Query Match 12.2%; Score 6; DB 2; Length 269;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 39 QEAYR 44  
Db 214 QEAYR 219

Search completed: December 4, 2002, 15:40:02  
Job time: 21 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:34:59 : Search time 11 Seconds  
(without alignments)  
184.758 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 49  
Sequence: 1 YLYQWLGAIPVPPDPLEPRR.....DELADHTGQEAAYRREYGPV 49

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	100	1 OSTC_HUMAN	P02818 homo sapien
2	39	79.6	49	1 OSTC_HORSE	P3005 equus caball
3	30	61.2	48	1 OSTC_DRONO	P15504 dromaelus no
4	30	61.2	49	1 OSTC_MACFA	P02819 macaca fasc
5	30	61.2	49	1 OSTC_XENLA	P40147 xenopus lae
6	30	61.2	100	1 OSTC_BOVIN	P02820 bos taurus
7	17	34.7	49	1 OSTC_FELCA	P02821 felis silve
8	15	30.6	49	1 OSTC_RABIT	P39056 oryctolagus
9	14	28.6	49	1 OSTC_CANFA	P1455 canis famli
10	14	28.6	97	1 OSTC_CHICK	P02822 gallus gall
11	11	22.4	99	1 OSTC_RAT	P04640 rattus norv
12	7	14.3	45	1 OSTC_SPAU	P40148 sparus aura
13	7	14.3	47	1 OSTC_XIPGL	P02823 xiphias gla
14	7	14.3	447	1 NRH3_HUMAN	Q13133 homo sapien
15	7	14.3	586	1 UL84_HCMVA	P29839 human cytom
16	7	14.3	587	1 UL84_HCMVA	P29839 human cytom
17	7	14.3	1153	1 JAK1_BRARE	O12990 brachydanio
18	6	12.2	70	1 RS21_HELPJ	Q92119 helicobacte
19	6	12.2	70	1 RS21_HELPJ	P56028 helicobacte
20	6	12.2	95	1 OSTC_MOUSE	P04641 mus musculu
21	6	12.2	95	1 OSTC_MOUSE	P04641 mus musculu
22	6	12.2	103	1 OSTC_MOUSE	P5615 mus musculu
23	6	12.2	103	1 OSTC_MOUSE	P5615 mus musculu
24	6	12.2	116	1 CYP2_COTJA	G92742 chlamydia p
25	6	12.2	119	1 MP2_MERDI	P81061 coturnix co
26	6	12.2	139	1 CYP2_CHICK	P80235 nereis dive
27	6	12.2	171	1 YKH6_GAEEL	P01038 gallus gall
28	6	12.2	210	1 ACUA_BACSU	P34274 caenorhabdi
29	6	12.2	304	1 NADA_MERTH	P39065 methanobact
30	6	12.2	338	1 YOJM_BACSU	O37855 methanobact
31	6	12.2	349	1 Y028_BORBU	P54550 bacillus su
32	6	12.2	367	1 Y797_MERJA	O51059 borrelia bu
33	6	12.2	399	1 B4G1_MOUSE	O58207 methanococc
				1 IRTF_MOUSE	P15535 m beta-1,4-
					O61179 mus musculu

## ALIGNMENTS

RESULT 1	OSTC_HUMAN	STANDARD	PRT	100 AA.
AC	P02818;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)			
DE	(Bone Gla-protein) (BGP).			
GN	BGLAP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE-90245603; PubMed-2336375;			
RA	Kiefer M.C., Saphire A.C.S., Bauer D.M., Barr P.J.;			
RT	"The cDNA and derived amino acid sequences of human and bovine bone			
RL	Nucleic Acids Res. 18:1909-1909(1990).			
RN	[2]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE-87004555; PubMed-3019668;			
RA	Celeste A.J., Buckner J.L., Kriz R., Wang E.A., Mooney J.M.;			
RT	"Isolation of the human gene for bone gla protein utilizing mouse and			
RL	rat cDNA clones.";			
EMBO	J. 5:1885-1890(1986).			
RN	[3]			
RX	SEQUENCE OF 52-100.			
RX	MEDLINE-8106914; PubMed-6967872;			
RA	Poser J.W., Esch F.S., Ling N.C., Price P.A.;			
RT	"Isolation and sequence of the vitamin K-dependent protein from human			
RL	bone. Undercarboxylation of the first glutamic acid residue.";			
J. Biol. Chem.	255:8685-8691(1980).			
CC	-1- FUNCTION: CONSTITUTES 1-28 OF THE TOTAL BONE PROTEIN. IT BINDS			
CC	STRONGLY TO APATITE AND CALCIUM.			
CC	-1- PPM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K			
CC	DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE			
CC	BINDING OF CALCIUM.			
CC	-1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN			
CC	FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X53698; CA37736.1; -			
DR	EMBL; X51699; CA35996.1; -			
DR	EMBL; X04143; CA27763.1; -			
DR	PIR; A03301; GEHU.			
DR	PIR; C25471; C25471.			

34	6	12.2	413	1	PURK_CORAM	Q44678 corynebacte
35	6	12.2	445	1	CLPX_ANASP	O89477 anabena sp
36	6	12.2	461	1	HMCS_ARATH	P54873 arabidopsis
37	6	12.2	471	1	STEL_TERPE	O82670 arabidopsis
38	6	12.2	532	1	ICPO_HSVB	P28990 equine herp
39	6	12.2	553	1	NOEB_RHIME	O52893 rhizobium m
40	6	12.2	725	1	ECHP_CAVPO	P55100 cavia porce
41	6	12.2	765	1	DPP4_BOVIN	P81425 bos taurus
42	6	12.2	788	1	TRSL_HCMVA	P09695 human cytom
43	6	12.2	806	1	PLSB_ECO57	P58130 escherichia
44	6	12.2	806	1	PLSB_ECOLI	P00482 escherichia
45	6	12.2	806	1	PLSB_SALTI	Q8216 salmonella

RESULT 4	OSTC_MACFA	STANDARD:	PRT:	49 AA.
ID	OSTC_MACFA			
AC	P02819;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).			
GN	BGLAP.			
OS	Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;			
OC	Cercopithecoinae; Macaca.			
NCBI	_taxid=9541;			
TX				

```

RN [1]
RP MEDLINE-82182842; PubMed-6978733;
RX Hauschka P.V., Carr S.A., Blemann K.;
RT "Primary structure of monkey osteocalcin.";
RL Biochemistry 21:638-642(1982).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
DR PIR: A03302; GEMKI
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;
BM Bone.
FT MOD_RES 9 9 HYDROXYLATION.
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISUFID 23 29 BY SIMILARITY.
SQ SEQUENCE 49 AA; 5743 MW; C20116014D0C4958 CRC64;

Query Match 61.2%; Score 30; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 1;le-26;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 REVELNPDCELDADHIGFOEAYRRFGPV 49
DB 20 REVELNPDCELDADHIGFOEAYRRFGPV 49

RESULT 5
OSTC_XENLA
ID OSTC_XENLA STANDARD; PRT; 49 AA.
AC PA0147;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
OX NCBI_Taxid=8335;
RN [1]
RP SEQUENCE.
RX Cancela M., Williamson M.K., Price P.A.;
RT "Amino-acid sequence of bone Gla protein from the African clawed toad
RT Xenopus laevis and the fish Sparus aurata.";
RL Int. J. Pept. Protein Res. 46:419-423(1995).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone.

```

```

FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISUFID 23 29 BY SIMILARITY.
SQ SEQUENCE 49 AA; 5360 MW; 7A9ABF63A12E6047 CRC64;

Query Match 61.2%; Score 30; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 1;le-26;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 REVELNPDCELDADHIGFOEAYRRFGPV 49
DB 20 REVELNPDCELDADHIGFOEAYRRFGPV 49

RESULT 6
OSTC_BOVIN
ID OSTC_BOVIN STANDARD; PRT; 100 AA.
AC P02820;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)
DE (Bone Gla-protein) (BGP).
GN BGLAP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90245603; PubMed-2336375;
RA Kiefer M.C., Sapphire A.C.S., Bauer D.M., Barr P.J.;
RT "The cDNA and derived amino acid sequences of human and bovine bone
RT Gla protein.";
RL Nucleic Acids Res. 18:1909-1909(1990).
RN [2]
RP SEQUENCE OF 52-100.
RX MEDLINE-77036749; PubMed-1068450;
RA Price P.A., Poser J.W., Raman N.;
RT "Primary structure of the gamma-carboxyglutamic acid-containing
RT protein from bovine bone.";
RL Proc. Natl. Acad. Sci. U.S.A. 73:3374-3375(1976).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53699; CAA37737.1; -
DR EMBL: X51700; CAA35997.1; -
DR PIR: A03303; GEMO
DR PIR: S12653; S12653.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
KW Bone; Signal.
FT SIGNAL 1 23

```

FT PROPEP 24 51 PROBABLE.  
 FT CHAIN 52 100 OSTEOCALCIN.  
 FT MOD\_RES 60 60 HYDROXYLATION.  
 FT MOD\_RES 68 68 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 72 72 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 75 75 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 74 80  
 SQ SEQUENCE 100 AA: 11042 MW: 73015D1681B26219 CRC64;

Query Match 61.2%; Score 30; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-26;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 REYCELNPCDELADHIGFEAYRRFGPV 49  
 DB 71 REYCELNPCDELADHIGFEAYRRFGPV 100

RESULT 7  
 OSTC\_FELCA STANDARD; PRT; 49 AA.  
 AC P02821;

DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).

OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxId=9685;

RP SEQUENCE.  
 RX MEDLINE=85054706; PubMed=6334077;  
 RA Shimomura H., Kanai Y., Sanada K.;  
 RT "Primary structure of cat osteocalcin.";  
 RL J. Biochem. 96:405-411(1984).

CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN FAMILY.

DR PIR: A03304; GECF.  
 DR InterPro: IPR002384; GLA bone.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS; PR00002; GLABONE.

DR SMART; SM00069; GLA; 1.  
 DR PROSITE; PS00011; GLT\_CARBOXYLATION; 1.  
 KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;  
 KW Bone.

FT MOD\_RES 9 9 HYDROXYLATION.  
 FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 23 29  
 SQ SEQUENCE 49 AA: 5495 MW: 93D2131FA9F656D3 CRC64;

Query Match 34.7%; Score 17; DB 1; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 2,9e-12;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 CEINPCDELADHIGFQ 39  
 DB 23 CEINPCDELADHIGFQ 39

RESULT 8  
 OSTC\_RABIT STANDARD; PRT; 49 AA.  
 ID OSTC\_RABIT

AC P39056;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).

OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxId=9986;

RP SEQUENCE.  
 RX MEDLINE=92175242; PubMed=1794506;  
 RA Viridi A.S., Willis A.C., Hauschka P.V., Triffitt J.T.;  
 RT "Primary aminoacid sequence of rabbit osteocalcin.";  
 RL Biochem. Soc. Trans. 19:3735-3738(1991).

CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE BINDING OF CALCIUM.

CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN FAMILY.  
 DR PIR: A61280; A61280.  
 DR InterPro: IPR002384; GLA bone.

DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS; PR00002; GLABONE.

DR SMART; SM00069; GLA; 1.  
 DR PROSITE; PS00011; GLT\_CARBOXYLATION; 1.  
 KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;  
 KW Bone.

FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 9 9 HYDROXYLATION.  
 FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 23 29 BY SIMILARITY.  
 SQ SEQUENCE 49 AA: 5431 MW: 7B218871F0312253 CRC64;

Query Match 30.6%; Score 15; DB 1; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 4,8e-10;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 REYCELNPCDELAD 34  
 DB 20 REYCELNPCDELAD 34

RESULT 9  
 OSTC\_CANFA STANDARD; PRT; 49 AA.  
 ID OSTC\_CANFA

AC P81455;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-protein) (BGP).

OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxId=9615;

RP SEQUENCE.  
 RX MEDLINE=93318657; PubMed=8101026;  
 RA Colombo G., Fantl P., Yao C., Malluche H.H.;  
 RT "Isolation and complete amino acid sequence of osteocalcin from canine bone.";

RL J. Bone Miner. Res. 8:733-743(1993).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS STRONGLY TO APATITE AND CALCIUM.



```
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;
  Bone.
FT MOD_RES 9 9 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 23 29
SQ SEQUENCE 49 AA; 5524 MW; 43121D015817CEA6 CRC64;

Query Match 28.6%; Score 14; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 6; 2e-09;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CDELAHDHIGFOEAY 42
Db 29 CDELAHDHIGFOEAY 42

RESULT 10
OSTC_CHICK STANDARD; PRT; 97 AA.
AC P02822; O90620;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)
DE (Bone Gla-protein) (BGP).
GN BGLAP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Bone;
RX MEDLINE=9526465; PubMed=7747623;
RA Neugebauer B.M.; Moore M.A.; Broess M.; Gerstenfeld L.C.;
RA Hauschka P.V.;
RT "Characterization of structural sequences in the chicken osteocalcin
RT gene: expression of osteocalcin by maturing osteoblasts and by
RT hypertrophic chondrocytes in vitro.";
RL J. Bone Miner. Res. 10:157-163(1995).
RN [2]
RP SEQUENCE OF 49-97.
RX MEDLINE=82007831; PubMed=6792200;
RA Carr S.A.; Hauschka P.V.; Bleemann K.;
RT "Gas chromatographic mass spectrometric sequence determination of
RT osteocalcin, a gamma-carboxyglutamic acid-containing protein from
RT chicken bone.";
RL J. Biol. Chem. 256:9944-9950(1981).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U10578; AAT78809.1; -.
DR PIR: A03305; GECH.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone; Signal.
FT SIGNAL 1 20
FT PROPEP 21 48 OSTEOCALCIN.
FT MOD_RES 49 97 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 65 65 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 72 72 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 71 77 BY SIMILARITY.
SQ SEQUENCE 97 AA; 10707 MW; 768E8685C3327D62 CRC64;

Query Match 28.6%; Score 14; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 1; 1e-08;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IGFOEAYRRFYGPV 49
Db 84 IGFOEAYRRFYGPV 97

RESULT 11
OSTC_RAT STANDARD; PRT; 99 AA.
AC P04640;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)
DE (Bone Gla-protein) (BGP).
GN BGLAP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87004555; PubMed=3019668;
RA Celeste A.J.; Buecker J.L.; Kritz R.; Wang E.A.; Moxney J.M.;
RT "Isolation of the human gene for bone gla protein utilizing mouse and
RT rat cDNA clones.";
RL EMBO J. 5:1885-1890(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85298305; PubMed=3875856;
RA Pan L.C.; Price P.A.;
RT "The propeptide of rat bone gamma-carboxyglutamic acid protein shares
RT homology with other vitamin K-dependent protein precursors.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6109-6113(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89118266; PubMed=3265336;
RA Yoon K.; Rutledge S.J.C.; Buena R.F.; Rodan G.A.;
RT "Characterization of the rat osteocalcin gene: stimulation of
RT promoter activity by 1,25-dihydroxyvitamin D3.";
RL Biochemistry 27:8521-8526(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89251082; PubMed=2785907;
RA Theobald G.; Haberstroch L.M.; Price P.A.;
RT "Molecular structure of the rat bone Gla protein gene and
RT identification of putative regulatory elements.";
```

```

RL DNA 8:213-221(1989).
RN [5]
RN SEQUENCE FROM N.A.
RX MEDLINE=89145200; PubMed=2784002;
RA Lian J., Stewart C., Puchacz E., Mackowiak S., Shalhoub V.,
RA Collart D., Zambetti G., Stein G.,
RT "Structure of the rat osteocalcin gene and regulation of vitamin D-
RT dependent expression."
RL Proc. Natl. Acad. Sci. U.S.A. 86:1143-1147(1989).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTH: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X04141; CAA27761.1; -
DR EMBL: M11777; AAA40816.1; -
DR EMBL: M23637; AAA41761.1; -
DR EMBL: M25490; AAA53280.1; -
DR EMBL: J04500; AAA41764.1; -
DR PIR: A25167; GERT.
DR PIR: A25471; A25471.
DR PIR: A31856; A31856.
DR PIR: A31419; A31419.
DR PIR: A32324; A32324.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR SMART: SM00069; GLA; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
KW Calcium-binding; gamma-carboxyglutamic acid; vitamin K; Hydroxylation;
KW Bone; Signal.
FT SIGNAL 1 23 PROBABLE.
FT PROPEP 24 49 PROBABLE.
FT CHAIN 50 99 OSTEOCALCIN.
FT MOD_RES 58 58 HYDROXYLATION.
FT MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID
FT MOD_RES 78 78 GAMMA-CARBOXYGLUTAMIC ACID
FT DISULFID 72 78 BY SIMILARITY.
FT SEQUENCE 99 AA; 10927 MW; 7F18F166D4E4388 CRC64;
Query Match 22.4%; Score 11; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 CDELDHIGFQ 39
DB 78 CDELDHIGFQ 88
RESULT 12
OSTC_SPAU STANDARD; PRT; 45 AA.
AC P40148;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-

```

```

DE protein) (BGP).
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RN SEQUENCE.
RX MEDLINE=96139691; PubMed=8567186;
RA Cancela M., Williamson M.K., Price P.A.;
RT "Amino-acid sequence of bone Gla protein from the African clawed toad
RT Xenopus laevis and the fish Sparus aurata."
RL Int. J. Pept. Protein Res. 46:419-423(1995).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTH: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
CC -----
CC InterPro: IPR002384; GLA_bone.
CC InterPro: IPR000294; VitK_dep_GLA.
CC Pfam: PF00594; gla; 1.
CC PRINTS: PR00002; GLABONE.
CC PROSITE: PS00011; GLU-CARBOXYLATION; PARTIAL.
KW Calcium-binding; gamma-carboxyglutamic acid; vitamin K; Bone.
FT MOD_RES 11 11 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 15 15 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 18 18 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 17 23 BY SIMILARITY.
FT SEQUENCE 45 AA; 4951 MW; 5BAEB6A1DCB4A18 CRC64;
Query Match 14.3%; Score 7; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 REVCEIN 26
DB 14 REVCEIN 20
RESULT 13
OSTC_XIPGL STANDARD; PRT; 47 AA.
AC P02823;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
OS Xiphias gladius (Swordfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Xiphiidae; Xiphias.
OX NCBI_TaxID=8245;
RN [1]
RN SEQUENCE.
RX Price P.A., Otsuka A.S., Poser J.W.;
RT "Comparison of gamma-carboxyglutamic acid-containing proteins from
RT bovine and swordfish bone: primary structure and Ca++ binding."
RL (In) Wasserman R.H., Corradino R.A., Carafoli E., Kretsinger R.H.,
RL MacLennan D.H., Siegel F.L. (eds.);
RL Calcium-binding proteins and calcium function, pp.333-337,
RL Elsevier, New York (1977).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTH: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.

```

```

DR PIR: A03306; GEMF.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; Vltk_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00002; GLABONE.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
KW Calcium-binding; gamma-carboxyglutamic acid; Vitamin K; Bone.
FT MOD_RES 13 13 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 20 20 GAMMA-CARBOXYGLUTAMIC ACID.
FT DISULFID 19 25
SQ SEQUENCE 47 AA; 5080 MW; 8D93FE1787B09223 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 REVCELN 26
Db 16 REVCELN 22

RESULT 14
NRH3_HUMAN STANDARD; PRT; 447 AA.
AC 013133;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxytocin receptor LXR-alpha (Liver X receptor alpha) (Nuclear orphan
DE receptor LXR-alpha).
GN NRH3 OR LXR.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95262897; PubMed=7744246;
RA Willy P.J., Umeson K., Ong E.S., Evans R.M., Heyman R.A.,
RA Mangelsdorf D.J.;
RT "LXR, a nuclear receptor that defines a distinct retinoid response
RT pathway.";
RL Genes Dev. 9:1033-1045(1995).
CC -!- FUNCTION: ORPHAN RECEPTOR. INTERACTION WITH RXR SHIFTS RXR FROM
CC ITS ROLE AS A SILENT DNA-BINDING PARTNER TO AN ACTIVE LIGAND-
CC BINDING SUBUNIT IN MEDIATING RETINOID RESPONSES THROUGH TARGET
CC GENES DEFINED BY LXRES. LXRES ARE DR4-TYPE RESPONSE ELEMENTS
CC CHARACTERIZED BY DIRECT REPEATS OF TWO SIMILAR HEXANUCLEOTIDE HALF-
CC SITES SPACED BY FOUR NUCLEOTIDES. LXRA PLAYS AN IMPORTANT ROLE IN
CC THE REGULATION OF CHOLESTEROL HOMEOSTASIS.
CC -!- SUBUNIT: HETERODIMER OF LXRA AND RXR.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: VISCERAL ORGANS SPECIFIC EXPRESSION. STRONG
CC EXPRESSION WAS FOUND IN LIVER, KIDNEY AND INTESTINE FOLLOWED BY
CC SPLEEN AND TO A LESSER EXTENT THE ADRENALS.
CC -!- INDUCTION: BY 9-CIS RETINOIC ACID (9CRA).
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRL SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U22662; AA85856.1; -.
CC HSSP: P03372; IHCO.
CC TRANSFAC: T02752; -.
CC Genew: HGNC:7966; NRH3.

```

```

DR MIM: 602423; -.
DR InterPro: IPR000536; Hormone_rec_119.
DR InterPro: IPR001723; Stthrmn_receptor.
DR InterPro: IPR001628; ZnF_C4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR PRINTS: PR00398; SPRODHOMNER.
DR PRINTS: PR00047; SPROIDFORGER.
DR PRODOM: PD000035; ZnF_C4steroid; 1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; ZnF_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
FT DNA_BIND 98 163 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 98 118 C4-TYPE.
FT ZN_FING 134 158 C4-TYPE.
FT DOMAIN 215 434 LIGAND-BINDING (POTENTIAL).
SQ SEQUENCE 447 AA; 50481 MW; A227B233DB0F8C96 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WLGAVP 11
Db 4 WLGAVP 10

RESULT 15
UL84_HCMVA STANDARD; PRT; 586 AA.
AC P16727;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE 65 kDa early nonstructural protein (UL84 protein).
GN UL84.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA Peddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X17403; CA35358.1; -.
CC PIR: S09848; KMBEDE.
CC Nonstructural protein.
KW DOMAIN 9 19 ARG-RICH (BASIC).
KW DOMAIN 162 170 LYS-RICH (BASIC).
FT DOMAIN 171 182 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 586 AA; 65428 MW; 54AB912D6077223F CRC64;

Query Match 14.3%; Score 7; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 YRRTYCP 48

```

```

DB      443 YRRFYGP 449
      |||||
RESULT 16
ID      UL84_HCMVT      STANDARD;      PRT;      587 AA.
AC      P29839;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DE      65 kDa early nonstructural protein (UL84 protein).
GN      UL84.
OS      Human cytomegalovirus (strain Towne).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Betaherpesvirinae; Cytomegalovirus.
ON      NCBI_TaxId=10363;
RX      MEDLINE=92114132; PubMed=1309892;
RA      He Y.S., Xu L., Huang E.S.;
RT      Characterization of human cytomegalovirus UL84 early gene and
RT      identification of its putative protein product.
RL      J. Virol. 66:1098-1108(1992).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M81432; AAA5947.1; -.
DR      PIR; A41808; WMBETE.
KM      Nonstructural protein.
FT      DOMAIN 9 19 ARG-RICH (BASIC).
FT      DOMAIN 162 170 LYS-RICH (BASIC).
FT      DOMAIN 171 183 ASP/GLU-RICH (ACIDIC).
SQ      SEQUENCE 587 AA; 65388 MW; 13C170B41FB3220B CRC64;

Query Match      14.3%; Score 7; DB 1; Length 587;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      42 YRRFYGP 48
      |||||
DB      444 YRRFYGP 450

RESULT 17
ID      JAK1_BRARE      STANDARD;      PRT;      1153 AA.
AC      O12980; C73880;
DT      15-JUN-2002 (Rel. 41, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DE      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Tyrosine-protein kinase Jak1 (EC 2.7.1.112) (Janus kinase 1) (Jak-1).
GN      JAK1.
OS      Brachydanio rerio (Zebrafish) (Danio rerio).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_TaxId=7955;
RN      [1]
RP      SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, MUTAGENESIS OF LYS-905, AND
RP      FUNCTION.
RC      TISSUE=castrula;
RA      MEDLINE=97250493; PubMed=9096349;
RA      Conway G., Margoliath A., Wong-Madden S., Roberts R.J., Gilbert W.;
RT      *jak1 kinase is required for cell migrations and anterior
RT      specification in zebrafish embryos.
RT      Proc. Natl. Acad. Sci. U.S.A. 94:3082-3087(1997).

```

```

RN      [2]
RP      SEQUENCE OF 443-1153 FROM N.A.
RC      TISSUE=Embryo;
RX      MEDLINE=99445372; PubMed=10515866;
RA      Oates A.C., Brownlie A., Pratt S.J., Irvine D.V., Liao E.C., Paw B.H.,
RA      Dorian K.J., Johnson S.L., Postlethwait J.H., Zon L.I., Wilks A.F.;
RT      *gene duplication of zebrafish JAK2 homologs is accompanied by
RT      divergent embryonic expression patterns: only jak2a is expressed
RT      during erythropoiesis.
RL      Blood 94:2622-2636(1999).
CC      -1- FUNCTION: Tyrosine kinase of the non-receptor type. Appears to be
CC      required in early development for specific cell migrations
CC      (epiboly), expression of homeobox protein goosecoid and formation
CC      of anterior structures.
CC      -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC      tyrosine phosphate.
CC      -1- SUBCELLULAR LOCATION: Wholly intracellular, possibly membrane
CC      associated (By similarity).
CC      -1- DEVELOPMENTAL STAGE: Present in the unfertilized egg through to
CC      the blastula stage where it is distributed uniformly. Levels drop
CC      rapidly at four hours development, remain very low until 10 hours,
CC      then gradually increase from 12 hours with a rapid increase at 48
CC      hours. At 48 hours it is concentrated in the region of the gill
CC      arches. Also present in the adult.
CC      -1- DOMAIN: Possesses two phosphotransferase domains. The second one
CC      probably contains the catalytic domain, while the presence of
CC      slight differences suggest a different role for domain 1.
CC      -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. JAK
CC      SUBFAMILY.
CC      -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U82980; AAB54114.1; -.
DR      EMBL; AJ005689; CAA06673.1; -.
DR      HSSP; P11362; IFCK.
DR      ZFIN; ZDB-GENE-980526-142; jak1.
DR      InterPro: IPR000299; Band_4.1.
DR      InterPro: IPR000719; Euk_pkinase.
DR      InterPro: IPR000980; SH2.
DR      InterPro: IPR001245; Tyr_pkinase.
DR      Pfam; PF00017; SH2; 1.
DR      Pfam; PF00069; pkinase; 2.
DR      ProDom; PD000001; Euk_pkinase; 2.
DR      ProDom; PD000093; SH2; 1.
DR      SMART; SM00295; BA1; 1.
DR      SMART; SM00252; SH2; 1.
DR      SMART; SM00219; TYKc; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.
DR      PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KM      Transferase: Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KM      SH2 domain; Repeat; Developmental protein.
FT      DOMAIN 336 356 LYS-RICH (BASIC).
FT      DOMAIN 435 540 SH2 (ATYPICAL).
FT      DOMAIN 580 846 PROTEIN KINASE 1.
FT      DOMAIN 872 1150 PROTEIN KINASE 2.
FT      NP_BIND 878 886 ATP (BY SIMILARITY).
FT      BINDING 905 905 ATP.
FT      ACT_SITE 1000 1000 BY SIMILARITY.
FT      MOD_RES 1031 1031 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT      MUTAGEN 905 905 K->E: LOSS OF AUTOPHOSPHORYLATION AND
FT      DEFECTS IN EARLY DEVELOPMENT.
FT      CONFLICT 649 649 T -> I (IN REF. 2).
FT      CONFLICT 770 770 T -> S (IN REF. 2).
FT      CONFLICT 799 799 S -> T (IN REF. 2).
FT      CONFLICT 978 978 H -> L (IN REF. 2).

```

```

FT  CONFLICT 1111 1111 L -> I (IN REF. 2).
SQ  SEQUENCE 1153 AA; 132480 MW; 736D5263D03E7450 CRC64;
Query Match
Best Local Similarity 14.3%; Score 7; DB 1; Length 1153;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  27 PDCDELA 33
    |||||
Db  814 PDCDELA 820

RESULT 18
RS21_HELPJ STANDARD; PRT; 70 AA.
AC  09ZLR9;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  30S ribosomal protein S21.
GN  RPSU OR JHP0509.
OS  Helicobacter pylori J99 (Campylobacter pylori J99).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OX  NCBI_TaxID=85963;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99120557; PubMed=9923682;
RA  Alm R.A., Ling L.-S.L., Molt D.T., King B.L., Brown E.D., Dolg P.C.,
RA  Smith D.R., Noonan B., Guild B.C., deJonge B.L., Camel G.,
RA  Tummano P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA  Gibson R., Metberg D., Mills S.D., Jiang Q., Taylor D.E., Voyts G.F.,
RA  Trust T.J.;
RT  "Genomic sequence comparison of two unrelated isolates of the human
RT  gastric pathogen Helicobacter pylori."
RL  Nature 397:176-180(1999).
CC  -1- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: AE001484; AAD06085.1; -
CC  InterPro: IPR001911; Ribosomal_S21.
CC  Pfam: PF01165; Ribosomal_S21; 1.
CC  PRINTS: PR00976; RIBOSOMAL_S21.
CC  ProDom: PD005521; Ribosomal_S21; 1.
CC  TIGRfams: TIGR00030; S21P; 1.
CC  DR PROSITE: PS01181; RIBOSOMAL_S21; 1.
CC  DR PROSITE: PS01181; RIBOSOMAL_S21; 1.
CC  KW Ribosomal protein; Complete proteome.
CC  SQ SEQUENCE 70 AA; 8614 MW; B1DA8696B1A03B7B CRC64;

Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  40 EAYRRF 45
    |||||
Db  14 EAYRRF 19

RESULT 19
RS21_HELPJ STANDARD; PRT; 70 AA.
AC  P56028;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  30S ribosomal protein S21.

```

```

GN  RPSU OR HP0562.
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC  Helicobacter.
OX  NCBI_TaxID=210;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=26695 / ATCC 700392;
RX  MEDLINE=97394467; PubMed=9252185;
RA  Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA  Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA  Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
RA  Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA  McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA  Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA  Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA  Hayes W.S., Borodovsky M., Kar P.D., Smith H.O., Fraser C.M.,
RA  Venter J.C.;
RT  "The complete genome sequence of the gastric pathogen Helicobacter
RT  pylori."
RL  Nature 388:539-547(1997).
CC  -1- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: AE000570; AAD07628.1; -
CC  TIGR: HP0562; -
CC  DR InterPro: IPR001911; Ribosomal_S21.
CC  DR Pfam: PF01165; Ribosomal_S21; 1.
CC  PRINTS: PR00976; RIBOSOMAL_S21.
CC  ProDom: PD005521; Ribosomal_S21; 1.
CC  TIGRfams: TIGR00030; S21P; 1.
CC  DR PROSITE: PS01181; RIBOSOMAL_S21; 1.
CC  DR PROSITE: PS01181; RIBOSOMAL_S21; 1.
CC  KW Ribosomal protein; Complete proteome.
CC  SQ SEQUENCE 70 AA; 8613 MW; 417A8696B1A03B76 CRC64;

Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  40 EAYRRF 45
    |||||
Db  14 EAYRRF 19

RESULT 20
OSTC_MOUSE STANDARD; PRT; 95 AA.
ID  OSTC_MOUSE
AC  P04641;
DT  13-AUG-1987 (Rel. 05, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)
DE  (Bone gla-protein) (BCP).
GN  BGLAP1 AND BGLAP2.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=87004555; PubMed=3019668;
RA  Celeste A.J., Buecker J.L., Kriz R., Wang E.A., Wozney J.M.,
RA  "Isolation of the human gene for bone gla protein utilizing mouse and
RA  rat cDNA clones."
RL  EMBO J. 5:1885-1890(1986).
RN  [2]

```

RP SEQUENCE FROM N.A.  
 RX MEDLINE=94062692; PubMed=8243336;  
 RA Rahman S., Oberdorf A., Montecino M., Tanhauser S.M., Lian J.B.,  
 RA Stein G.S., Laipis P.J., Stein J.L.;  
 RT "Multiple copies of the bone-specific osteocalcin gene in mouse and  
 rat.";  
 RL Endocrinology 133:3050-3053(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94117426; PubMed=8288580;  
 RA Despois C., Hogue D.A., Karsenty G.;  
 RT "The mouse osteocalcin gene cluster contains three genes with two  
 separate spatial and temporal patterns of expression.";  
 RL J. Biol. Chem. 269:1183-1190(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Calvaria;  
 RA Yotlov W.V., St Arnaud R.;  
 RT Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- TISSUE SPECIFICITY: BONE.  
 CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM  
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES CODING FOR OSTEOCALCIN, THEIR  
 CC CODING SEQUENCE ONLY DIFFERS IN THE SIGNAL PEPTIDE REGION.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: X04142; CAA27762.1; -;  
 DR EMBL: S67455; AAB29145.1; -;  
 DR EMBL: L24429; AAB39854.1; -;  
 DR EMBL: L24431; AAB39856.1; -;  
 DR EMBL: U11542; AAB40035.1; -;  
 DR PIR: B25471; B25471.  
 DR MGD: MGI:88156; Bglap1.  
 DR MGD: MGI:88157; Bglap2.  
 DR InterPro: IPR002384; VltK\_bone.  
 DR InterPro: IPR002394; VltK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PRO0002; GLABONE.  
 DR SMART: SM00069; GLA; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR KX Calciun-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone; Signal;  
 DR KX Multigene family.  
 FT SIGNAL 1 23 PROBABLE.  
 FT PROPEP 24 49  
 FT CHAIN 50 95 OSTEOCALCIN.  
 FT MOD\_RES 62 62 GAMMA-CARBOXYGLUTAMIC ACID  
 FT MOD\_RES 66 66 (BY SIMILARITY).  
 FT MOD\_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID  
 FT MOD\_RES 66 66 (BY SIMILARITY).  
 FT MOD\_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID  
 FT MOD\_RES 69 69 (BY SIMILARITY).  
 FT DISULFID 68 74 LS -> IF (IN BGLAP2).  
 FT VARIANT 4 5 A -> T (IN BGLAP2).  
 FT CONFLICT 11 11 A -> P (IN REF. 1).  
 FT CONFLICT 23 23 A -> P (IN REF. 1).  
 SQ SEQUENCE 95 AA: 10459 MW: D4AA611134805D9B CRC64;

Query Match 12.2%; Score 6; DB 1; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 8.2;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 PDPLER 18  
 DB 58 PDPLER 63  
 RESULT 21  
 ID OSTR\_MOUSE STANDARD; PRT; 95 AA.  
 AC P54615;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin-related protein precursor (CC-X) (Nephrocalcin).  
 GN BGLAP-RS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94117426; PubMed=8288580;  
 RA Despois C., Hogue D.A., Karsenty G.;  
 RT "The mouse osteocalcin gene cluster contains three genes with two  
 RT separate spatial and temporal patterns of expression.";  
 RL J. Biol. Chem. 269:1183-1190(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94062692; PubMed=8243336;  
 RA Rahman S., Oberdorf A., Montecino M., Tanhauser S.M., Lian J.B.,  
 RA Stein G.S., Laipis P.J., Stein J.L.;  
 RT "Multiple copies of the bone-specific osteocalcin gene in mouse and  
 rat.";  
 RL Endocrinology 133:3050-3053(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvJ; TISSUE=Liver;  
 RA Yotlov W.V., St Arnaud R.;  
 RT Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY BUT NOT IN BONE.  
 CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: L24430; AAB39855.1; -;  
 DR EMBL: S67456; AAB29146.1; -;  
 DR EMBL: U11541; AAB60445.1; -;  
 DR MGD: MGI:88155; Bglap-rs1.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR002394; VltK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PRO0002; GLABONE.  
 DR SMART: SM00069; GLA; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR KX Calciun-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone; Signal;  
 DR KX Multigene family.  
 FT SIGNAL 1 23 PROBABLE.  
 FT PROPEP 24 49  
 FT CHAIN 50 95 OSTEOCALCIN-RELATED PROTEIN.  
 FT MOD\_RES 62 62 GAMMA-CARBOXYGLUTAMIC ACID  
 FT MOD\_RES 66 66 (BY SIMILARITY).  
 FT MOD\_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID

```

FT      MOD_RES      69      69      (BY SIMILARITY).
FT      DISULFID      68      74      GAMMA-CARBOXYGLUTAMIC ACID
FT      SEQUENCE      95 AA; 10459 MW; 231968AA8B5848E CRC64;
SQ      (BY SIMILARITY).
Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      13 PDPLEP 18
        |||||
Db       58 PDPLEP 63

RESULT 22
Y592_CHLPN STANDARD: PRT: 103 AA.
AC 0927W2: 09K2D2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein CPN0592/CP0156/CPJ0592.
GN CPN0592 OR CP0156 OR CPJ0592.
OS Chlamydia pneumoniae (Chlamydothila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydothila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CM1029;
RX Keldman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Gilmwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heldelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RL pneumoniae AR39.";
RN [3]
RP Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hatake H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RL from Japan and CWL029 from USA.";
RN Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0161 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
CC EMBL: AE001643; AAD18731.1; -
DR EMBL: AE002177; AAF38037.1; ALT_INIT.
DR EMBL: AP002547; BAA98799.1; -
DR TIGR: CP0156; -
DR InterPro: IPR002696; DUF37.
DR Pfam: PF01809; DUF37; 1.
DR Prodom: PD004225; DUF37; 1.

```

```

DR      TIGR00278; DUF37; 1.
KW Hypothetical protein; Complete proteome.
SQ      SEQUENCE 103 AA; 11751 MW; CE9CA4852EA15A7C CRC64;
Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 103;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 YLYQWL 6
        |||||
Db       20 YLYQWL 25

RESULT 23
CYT_COTJA STANDARD: PRT: 116 AA.
AC P81061;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cystatin (Egg-white cystatin).
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE.
RC TISSUE=Egg white;
RX MEDLINE=97420480; PubMed=9276465;
RA Gernartz B., Engh R.A., Mentele R., Eckerskorn C., Torguato R.,
RA Wiltman J., Kolb H.J., Machleidt W., Ritz H., Auerwald E.A.;
RT "Quail cystatin: Isolation and characterisation of a new member of
RL the cystatin family and its hypothetical interaction with cathepsin
RT B.";
RN [1]
RP FEBS Lett. 412:551-558(1997).
CC -1- FUNCTION: THIS PROTEIN BINDS TIGHTLY TO AND INHIBITS PAPAIN AND
CC CATHEPSIN B.
CC -1- SIMILARITY: BELONGS TO THE CYSTATIN FAMILY.
CC -----
CC HSP: P01038; ICBM.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF000031; cystatin; 1.
DR SMART: SM00043; Cy; 1.
DR PROSITE: PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Phosphorylation.
FT ACT_SITE 9 9 REACTIVE SITE.
FT SITE 53 57 SECONDARY AREA OF CONTACT.
FT DISULFID 71 81
FT DISULFID 95 115
FT MOD_RES 80 80
SQ SEQUENCE 116 AA; 13093 MW; 48248621053A2F70 CRC64;
PHOSPHORYLATION.
Query Match
Best Local Similarity 12.2%; Score 6; DB 1; Length 116;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      6 LGAPPV 11
        |||||
Db       8 LGAPPV 13

RESULT 24
MP2_NERDI STANDARD: PRT: 119 AA.
AC P80255;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Non-metallothionein cadmium-binding protein (CD-BP) (MP II).
OS Nereis diversicolor (Sandworm) (Hediste diversicolor).
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
OC Phyllocorida; Nereididae; Nereidae; Nereididae; Nereididae;
OX NCBI_TaxID=6352;

```

RN [1]  
 RN SEQUENCE.  
 RX MEDLINE=94039032; PubMed=8223553;  
 RA Demyunck S., Li K.W., van der Schors R., Dhainaut-Courtois N.;  
 RT "Amino acid sequence of the small cadmium-binding protein (MP II)  
 RT from Nereis diversicolor (annelida, polychaeta). Evidence for a  
 RT myohemerythrin structure.";  
 RL Eur. J. Biochem. 217:151-156(1993).  
 RN [2]  
 RP SEQUENCE OF 1-33.  
 RX MEDLINE=91347123; PubMed=1908740;  
 RA Demyunck S., Sautiere P., van Beeumen J., Dhainaut-Courtois N.;  
 RT "Homologies between hemerythrins of sipunculids and cadmium-binding  
 RT metalloprotein (MP II) from a polychaete annelid, Nereis  
 RT diversicolor.";  
 RL C. R. Acad. Sci., III, Sci. Vie 312:317-322(1991).  
 CC -1- FUNCTION: MAY ACT AS A BUFFER TO CONTROL THE CONCENTRATION AND  
 CC THEREFORE THE TOXICITY OF CADMIUM.  
 CC -1- SIMILARITY: BELONGS TO THE HEMERYTHRIN FAMILY.  
 CC PIR: S38261; S38261.  
 DR HSSP: P02247; 2MR.  
 DR InterPro: IPR002063; Hemerythrin.  
 DR Pfam: PF01814; Hemerythrin; 1.  
 DR PRINTS: PR00186; HEMERYTHRIN.  
 DR ProDom: PD006099; Hemerythrin; 1.  
 DR TIGRfams: TIGR00058; Hemerythrin; 1.  
 DR PROSITE: PS00550; HEMERYTHRINS; 1.  
 KW Metal-binding; Cadmium; Cadmium resistance.  
 SQ SEQUENCE 119 AA; 13454 MW; F78F8DCE7F220E02 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 LGAPVP 11  
 |||||  
 Db 88 LGAPVP 93

## RESULT 25

CYT\_CHICK STANDARD: PRT; 139 AA.

AC P01038;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cystatin precursor (Egg-white cystatin).  
 OS Gallus gallus (Chicken).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 CC Gallus.  
 CC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90008873; PubMed=2793849;  
 RA Colella R., Sakaguchi Y., Nagase H., Bird J.W.C.;  
 RT "Chicken egg white cystatin. Molecular cloning, nucleotide sequence,  
 RT and tissue distribution.";  
 RL J. Biol. Chem. 264:17164-17169(1989).  
 RN [2]  
 RP SEQUENCE OF 24-139.  
 RX MEDLINE=84178305; PubMed=6712597;  
 RA Schwabe C., Anastasi A., Crow H., McDonald J.K., Barrett A.J.;  
 RT "Cystatin, amino acid sequence and possible secondary structure.";  
 RL Biochem. J. 217:813-817(1984).  
 RN [3]  
 RP SEQUENCE OF 24-139.  
 RX MEDLINE=84110059; PubMed=6662498;  
 RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,  
 RA Machleidt W.;  
 RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence  
 RT of cystatin from chicken egg white.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).

RN [4]  
 RP CHARACTERIZATION OF PROTEIN.  
 RX MEDLINE=83256421; PubMed=6409085;  
 RA Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J.H., Sayers C.A.,  
 RA Sunter D.C., Barrett A.J.;  
 RT "Cystatin, a protein inhibitor of cysteine proteinases. Improved  
 RT purification from egg white, characterization, and detection in  
 RT chicken serum.";  
 RL Biochem. J. 211:129-138(1983).  
 RN [5]  
 RP DISULFIDE BONDS.  
 RA Grubb A., Loeferberg H., Barrett A.J.;  
 RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken  
 RT cystatin.";  
 RL FEBS Lett. 170:370-374(1984).  
 RN [6]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=89252033; PubMed=2721673;  
 RA Lader B., Krieglstein K., Henschen A., Kos J., Turk V., Huber R.,  
 RA Bode W.;  
 RT "The cysteine proteinase inhibitor chicken cystatin is a  
 RT phosphoprotein.";  
 RL FEBS Lett. 248:162-168(1989).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=89052676; PubMed=3191914;  
 RA Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A.,  
 RA Brzin J., Kos J., Turk V.;  
 RT "The 2.0 A x-ray crystal structure of chicken egg white cystatin and  
 RT its possible mode of interaction with cysteine proteinases.";  
 RL EMBO J. 7:2593-2599(1988).  
 RN [8]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=94087719; PubMed=8263912;  
 RA Dieckmann T., Mitschang L., Hofmann M., Kos J., Turk V.,  
 RA Auerwald E.A., Jeanlck R., Oschkinat H.;  
 RT "The structures of native phosphorylated chicken cystatin and of a  
 RT recombinant unphosphorylated variant in solution.";  
 RL J. Mol. Biol. 234:1048-1059(1993).  
 CC -1- FUNCTION: THIS PROTEIN BINDS TIGHTLY TO AND INHIBITS A VARIETY OF  
 CC THIOLE PROTEASES INCLUDING FICIN, PAPAIN, AND CATHEPSINS B, C, H,  
 CC AND L. ALTHOUGH ISOLATED FROM EGG WHITE, IT IS ALSO PRESENT IN  
 CC SERUM.  
 CC -1- SIMILARITY: BELONGS TO THE CYSTATIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: J05077; AAA48744.1; -.  
 DR PIR: A01274; UDCH.  
 DR PIR: A34456; A34456.  
 DR PDB: ICW; 31-JAN-94.  
 DR PDB: 1A67; 27-MAY-98.  
 DR PDB: 1A90; 17-JUN-98.  
 DR InterPro: IPR000010; Cystatin.  
 DR Pfam: PF00031; cystatin; 1.  
 DR SMART: SM00043; CY; 1.  
 DR PROSITE: PS00287; CYSTATIN; 1.  
 KW Thiol protease inhibitor; Phosphorylation; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 139  
 FT ACT\_SITE 32 32  
 FT SITE 76 80  
 FT DISULFID 94 104  
 FT DISULFID 118 138  
 FT MOD\_RES 103 103  
 FT STRAND 35 36  
 FT TURN 39 40  
 PHOSPHORYLATION (PARTIAL).



```
FT HELIX 42 51
FT TURN 52 52
FT HELIX 53 56
FT TURN 57 58
FT STRAND 63 77
FT TURN 81 95
FT TURN 96 97
FT TURN 99 100
FT HELIX 101 108
FT STRAND 115 125
FT TURN 126 129
FT STRAND 130 138
SQ SEQUENCE 139 AA; 15287 MW; D92D1131C4D37891 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11
    111111
Db 31 LGAPVP 36

RESULT 26
YKH6_CAEEL STANDARD; PRT; 171 AA.
ID YKH6_CAEEL
AC P34274;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C02C2.6 in chromosome III.
GN C02C2.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Rhabdilitida; Rhabdilitidae;
OC Rhabdilitidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton L., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L23649; AAA27907.1; -
DR PIR: S44737; S44737.
DR MOPREP: C02C2.6; CE00032.
KW Hypothetical protein.
SQ SEQUENCE 171 AA; 20083 MW; 6EB3DFD3EDBD9B5 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 30 DELADH 35
    111111
Db 109 DELADH 114

RESULT 27
ACUA_BACSU STANDARD; PRT; 210 AA.
ID ACUA_BACSU
AC P39065;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acetoin utilization protein acua (EC 2.3.1.-).
GN ACUA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95020526; PubMed=7934817;
RA Grundy F.J., Waters D.A., Takova T.Y., Henkin T.M.;
RT "Identification of genes involved in utilization of acetate and
RT acetoin in Bacillus subtilis."
RL Mol. Microbiol. 10:259-271(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidis A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb trnB-dnaB region."
RL Microbiology 143:3431-3441(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98044033; PubMed=9384377;
RC STRAIN=168;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capiano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denliot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Ettian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karanata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidis A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potworlik S., Prescott A.M.,
RA Presecan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche M., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiuchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takenchi M., Takemochi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassatoli A.,
RA Viari A., Wambolt R., Wedler E., Wedler H., Weltergiger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -!- FUNCTION: ROLE IN GROWTH AND SPORULATION ON ACETOIN OR BUTANEDIOL.
CC INVOLVED IN THE BREAKDOWN OF THESE COMPOUNDS USED AS A CARBON
CC SOURCE. COULD ACT AS AN ACETYLTRANSFERASE.
CC -!- SIMILARITY: LOCAL. TO A VARIETY OF ACETYLTRANSFERASES.
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L17309; AAA68286.1; -  
DR EMBL: AF008220; AAC00396.1; -  
DR EMBL: Z99119; CAB14947.1; -  
DR PIR: S39645; S39645.  
DR Subtilisin; BG10369; acua.  
DR InterPro: IPR000182; GCN5acetyltransf.  
DR Pfam: PF00583; Acetyltransf. 1.  
DR Spoolation: Transferrase; Acyltransferase; Complete proteome.  
SQ SEQUENCE 210 AA; 24333 MW; 74DB94A465856DA CRC64;

Query Match 12.2%; Score 6; DB 1; Length 210;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 YPDPLE 17  
|11111  
DB 80 YPDPLE 85

RESULT 28  
NADA\_METTH STANDARD; PRT; 304 AA.  
ID NADA\_METTH  
AC 02785;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Guinoline synthetase A.  
GN NADA OR MTH1827.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales;  
OC Methanobacteriaceae; Methanobacter.  
OX NCBI\_TaxID=187420;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Delta H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
RA Aldege T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
RA Spadafora R., Viare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jiwani N., Caruso A., Bush D., Sater H., Patwell D., Prabhakar S.,  
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
RT delta: functional analysis and comparative genomics.";  
RL J. Bacteriol. 179:7135-7155(1997).

CC -1- FUNCTION: Catalyzes the condensation of iminosuccinate with  
CC dihydroxyacetone phosphate to form quinolinate (By similarity).  
CC -1- PATHWAY: NAD biosynthesis; aspartate to NAD; second step.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE QUINOLINATE SYNTHETASE A FAMILY.  
CC SUBFAMILY 2.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE000936; AAB86293.1; -  
DR InterPro: IPR003473; NADA.  
DR Pfam: PF02445; NADA; 1.  
DR TIGRfams: TIGR00550; nada; 1.

KW Pyridine nucleotide biosynthesis; Complete proteome.  
SQ SEQUENCE 304 AA; 34393 MW; B5C48ACE482143DD CRC64;

Query Match 12.2%; Score 6; DB 1; Length 304;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 ELADHI 36  
|11111  
DB 208 ELADHI 213

RESULT 29  
YQJM\_BACSU STANDARD; PRT; 338 AA.  
ID YQJM\_BACSU  
AC P54550.  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable NADH-dependent flavin oxidoreductase yqjM (EC 1.-.-.-).  
GN YQJM.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / JH642;  
RX MEDLINE=97124195; PubMed=8969508;  
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,  
RA Kobayashi Y.;  
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of  
RT the Bacillus subtilis genome containing the skin element and many  
RT sporulation genes.";  
RL Microbiology 142:3103-3111(1996).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D84432; BAA12619.1; -.
DR EMBL: Z99116; CAB14314.1; -.
DR Subtilisin; BG11742; yqjM.
DR InterPro: IPR001155; Oxidored_FMN.
DR Pfam: PF00724; Oxidored_FMN; 1.
DR Hypothetical protein; Oxidoreductase; NAD; FAD; Flavoprotein;
KM Complete proteome.
SO SEQUENCE 338 AA; 37583 MW; BDC52D34236326FE CRC64;

Query Match 12.2%; Score 6; DB 1; Length 338;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ADHIGF 38
Db 229 ADHIGF 234

RESULT 30
Y028_BORBU STANDARD; PRT; 349 AA.
ID Y028_BORBU
AC 051059;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BB0028 precursor.
GN BB0028.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterlind T., Wathley L., McDonald L., Aftach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AEO01117; AAC66428.1; -.
DR TIGR: BB0028; -.
KM Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 29
FT CHAIN 30 349 HYPOTHETICAL PROTEIN BB0028.
SO SEQUENCE 349 AA; 39705 MW; 915B1D34A3214226 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 EVELIN 26

```

```

Db 111 EVELIN 116

RESULT 31
Y797_METJA STANDARD; PRT; 367 AA.
ID Y797_METJA
AC 058207;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0797.
GN MJ0797.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uitterlind T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts R.M., Hurst M.A., Kaine B.P., Bordovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO M.JANNASCHII MJ1507.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67524; AAB98792.1; -.
DR TIGR: MJ0797; -.
DR InterPro: IPR003838; DUF214.
DR Pfam: PF02687; DUF214; 1.
KM Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 18 38
FT TRANSMEM 239 259 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
SO SEQUENCE 367 AA; 41038 MW; 3CC882FF21ED3746 CRC64;

Query Match 12.2%; Score 6; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRPYG 47
Db 178 YRRPYG 183

RESULT 32
B4G1_MOUSE STANDARD; PRT; 399 AA.
ID B4G1_MOUSE
AC P15535;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-1,4-galactosyltransferase 1 (EC 2.4.1.-) (Beta-1,4-GalTase 1)
DE (Beta4gal-T1) (b4gal-T1) (UDP-galactose:beta-N-acetylglucosamine beta-
DE 1,4-galactosyltransferase 1) (UDP-gal:beta-GlcNAc beta-1,4-

```



```

RESULT 33
IRTF_MOUSE
ID IRTF_MOUSE STANDARD: PRT: 399 AA.
AC Q61179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Transcriptional regulator ISGF3 gamma subunit (IFN-alpha responsive
transcription factor subunit) (interferon stimulated gene factor 3
gamma) (ISGF3 p48 subunit) (ISGF-3 gamma).
DE ISGF3G.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP MEDLINE-97037063; PubMed-8882710;
RX Suhara W., Yoneyama M., Yonekawa H., Takashi F.;
RT Structure of mouse interferon stimulated gene factor 3 gamma (ISGF3
gamma/p48) cDNA and chromosomal localization of the gene.";
RL J. Biochem. 119:231-234(1996).
CC -1- FUNCTION: RESPONSIBLE FOR THE INITIAL STIMULATION OF INF-ALPHA-
RESPONSIVE GENES. IT RECOGNIZES AND BINDS TO THE INF-STIMULATED
RESPONSE ELEMENT, OR ISRE WITHIN THE REGULATORY SEQUENCES OF A
TARGET GENES. ISGF3 PLAYS A PRIMARY ROLE IN THE TRANSMISSION OF A
SIGNAL FROM THE CELL SURFACE TO THE NUCLEUS.
CC -1- SUBUNIT: ACTIVE ISGF3 CONSISTS OF TWO COMPONENTS, A REGULATORY
MULTISUBUNIT COMPLEX. ISGF3 ALPHA (COMPOSED OF STAT1 ALPHA/BETA,
AND STAT2, ALSO DESIGNATED P91/P84 AND P133 RESPECTIVELY) AND
ISGF3 GAMMA (OR P48). THEY ASSEMBLE WHEN THE ISGF3 ALPHA COMPONENT
IS PHOSPHORYLATED VIA INF ACTIVATION.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- INDICATION: BY IFN-ALPHA AND IFN-BETA. UPON STIMULATION THE
REGULATORY PHOSPHORYLATED ALPHA AND BETA SUBUNITS ASSEMBLE WITH
THE GAMMA SUBUNIT AND TRANSLOCATE FROM THE CYTOPLASM TO THE
NUCLEUS.
CC -1- SIMILARITY: BELONGS TO THE IRTF FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U51992; AAC52494.1; -
DR HSSP: P23906; 21RF.
DR MGD: MGI:107387; ISGF3g.
DR InterPro: IPR001346; IRTF.
DR Pfam: PF00605; IRTF. 1.
DR PRINTS: PR00267; INTERFREGCT.
DR PRODOM: PD002355; IRTF. 1.
DR SMART: SM00348; IRTF. 1.
DR PROSITE: PS00601; IRTF. 1.
DR Transcription regulation; DNA-binding; Activator; Nuclear protein;
KW Interferon induction.
FT DNA_BIND 11 112 TRYPTOPHAN PENTAD REPEAT.
FT DOMAIN 165 170 POLY-GLY
SQ SEQUENCE 399 AA; 44610 MW; 92F9481ACEB665BD CRC64;
Query Match 12.2%; Score 6; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 PDPLEP 18
DB 256 PDPLEP 261

```

```

RESULT 34
PURI_CORAM
ID PURI_CORAM STANDARD: PRT: 413 AA.
AC Q4678;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)
DE (AIR carboxylase) (AIRC).
GN PURI.
OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).
OC Bacteria; Actinobacteria; Actinobacteria (Class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_Taxid=1697;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 6872;
RC MEDLINE-96236888; PubMed-8998996;
RA Chung S.O., Lee J.H., Lee S.Y., Lee D.S.;
RT "Genomic organization of purK and purL in Brevibacterium ammoniagenes
ATCC 6872: purL locus provides a clue for genomic evolution.";
RL FEMS Microbiol. Lett. 137:265-268(1996).
CC -1- FUNCTION: POSSESSES AN ATPASE ACTIVITY THAT IS DEPENDENT ON THE
PRESENCE OF AIR (AMINOIMIDAZOLE RIBONUCLEOTIDE). THE ASSOCIATION
OF PURK AND PURE PRODUCES AN ENZYME COMPLEX CAPABLE OF CONVERTING
AIR TO CAIR EFFICIENTLY UNDER PHYSIOLOGICAL CONDITION
(BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 5-amino-1-(5'-phospho-D-riboseyl)imidazole-4-
carboxylate + 5-amino-1-(5'-phospho-D-riboseyl)imidazole + CO(2).
CC -1- PATHWAY: De novo purine biosynthesis; sixth step.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PURK / PURT FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X91189; CAA62598.1; -
DR HSSP: P09029; 1B6R.
DR InterPro: IPR003135; ATP-grasp.
DR Pfam: PF02222; ATP-grasp. 1.
DR TIGRFAMs: TIGR01161; purK. 1.
KW Purine biosynthesis; Lyase; Decarboxylase.
SQ SEQUENCE 413 AA; 44110 MW; F738B230D5D119A CRC64;
Query Match 12.2%; Score 6; DB 1; Length 413;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 LGAPVP 11
DB 132 LGAPVP 137

```

```

RESULT 35
CLPX_ANASP
ID CLPX_ANASP STANDARD: PRT: 445 AA.
AC Q8YQX7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpX.
GN CLPX OR CLPC OR ALC3684.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_Taxid=103690;
RN [1]
RP SEQUENCE FROM N.A.

```

RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
RT cyanobacterium *Anabaena* sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
CC -1- FUNCTION: ATP-dependent specifically component of the Clp protease.  
CC it directs the protease to specific substrates. Can perform  
CC chaperone functions in the absence of clpP (by similarity).  
CC -1- SUBUNIT: Heterodimer of clpP and clpX (by similarity).  
CC -1- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AP003593; BAB75383.1; -  
DR InterPro; IPR003593; AAA\_Atpase.  
DR InterPro; IPR003595; AAA\_Atpase\_Centr.  
DR InterPro; IPR004487; ClpX.  
DR Pfam; PF00004; AAA; 1.  
DR SMART; SM00382; AAA; 1.  
DR TIGRFAMs; TIGR00382; clpX; 1.  
DR Chaperone; ATP-binding; Zinc-finger; Complete proteome.  
FT ZN\_FING 10 35 C4-TYPE.  
FT NP\_BIND 141 148 ATP (POTENTIAL).  
SQ SEQUENCE 445 AA; 48939 MW; C1E021D7831FDBA2 CRC64;  
  
Query Match 12.2%; Score 6; DB 1; Length 445;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 39 QEAYR 44  
| | | | |  
DB 381 QEAYR 386  
  
RESULT 36  
HMS\_ARATH  
ID HMS\_ARATH STANDARD: PRT: 461 AA.  
AC P54873; Q9S707;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hydroxymethylglutaryl-CoA synthase (EC 4.1.3.5) (HMG-CoA synthase)  
DE (3-hydroxy-3-methylglutaryl)-CoA synthase A synthase).  
GN HMG OR MYA1 OR AT4G11820 OR T26M18.30.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Landsberg erecta;  
RX MEDLINE=96144274; PubMed=8566777;  
RA Montanari F., Guillon M., Karst F., Delrot S.;  
RT "Isolation and characterization of a cDNA encoding Arabidopsis  
RT thaliana 3-hydroxy-3-methylglutaryl-coenzyme A synthase.";  
RL Gene 167:197-201(1995).  
RN  
RP  
RP Connolly E.L., Iearns R.M.;  
RT "Post-transcriptional regulation of HMG-CoA synthase expression in  
RT Arabidopsis thaliana.";  
RL Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.

RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20083488; PubMed=10617198;  
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,  
RA Harris B., Ansoerje W., Brandt P., Grivell L.A., Rieger M.,  
RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,  
RA Kreits M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,  
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
RA Vos P., Hohensel J., Zimmermann W., Medler H., Ridley P.,  
RA Langham S.-A., McCullagh B., Bilham L., Roben J., Vandenbussche F.,  
RA Van der Schueren J., Grymonprez B., Chung Y.-J., Vandenbussche F.,  
RA Breeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
RA Weitzenecker T., Bothe G., Ransperger U., Hilbert H., Braun M.,  
RA Holzner E., Brandt A., Peters S., Van Stevenen M., Dikse W.,  
RA Moolman P., Klein Lankester R., Rose M., Hauf J., Koetter P.,  
RA Benneiser S., Hempel S., Feldpausch M., Lambert S., Van den Daele H.,  
RA De Keyser A., Buysnaert C., Gielen J., Villarroel R., De Clercq R.,  
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,  
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Reumann S.,  
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Heitz A.,  
RA Neumann S., Argitlou A., Vitale D., Liguori R., Pivaraudi E.,  
RA Maassenet O., Quigley F., Clabaud G., Mendenhall A., Felder R.,  
RA Schnabl S., Hiller R., Schmidt W., Lechman A., Aubourg S.,  
RA Schenck F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
RA Gibbons T., Weber N., Vandenbol M., Baiges M., Terol J., Torres A.,  
RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,  
RA Frishman D., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Zaccaria P., Hayes M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
RA Nelson M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
RA Stoneking T., Kalicki J., Graves T., Hatton G., Edwards J.,  
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,  
RA Du H., All J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,  
RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,  
RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,  
RA Chen E., Marra M., Marijnsen R., McCombie W.R.;  
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
RT thaliana.";  
RL Nature 402:769-777(1999).  
CC -1- FUNCTION: THIS ENZYME CONDENSES ACETYL-CoA WITH ACETOACETYL-CoA TO  
CC FORM HMG-CoA, WHICH IS THE SUBSTRATE FOR HMG-CoA REDUCTASE (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxy-3-methylglutaryl-CoA + CoA =  
CC acetyl-CoA + H(2)O + acetoacetyl-CoA.  
CC -1- PATHWAY: PRODUCTION OF MEVALONATE FROM HMG-CoA PRIOR TO THE  
CC SYNTHESIS OF STEROLS AND ISOPRENOIDS.  
CC -1- SIMILARITY: BELONGS TO THE HMG-CoA SYNTHASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X83882; CAA58763.1; -  
DR EMBL; U79160; AAD00297.1; -  
DR EMBL; U79161; AAD00298.1; -  
DR EMBL; AL078606; CAB44320.1; -  
DR EMBL; AL161532; CAB78225.1; -  
DR InterPro; IPR000590; HMG-CoA\_synt.

DR Pfam: PF01154; HMG\_COA\_synth; 1.  
 DR PROSITE: PS01226; HMG\_COA\_SYNTHASE; 1.  
 KM Lyase; Sterol biosynthesis.  
 FT ACT\_SITE 117 117 POTENTIAL.  
 FT CONFLICT 306 306 A -> S (IN REF. 1).  
 FT CONFLICT 342 342 K -> N (IN REF. 1).  
 SQ SEQUENCE 461 AA: 51094 MW: F44908369AFB358B CRC64;

Query Match 12.2%; Score 6; DB 1; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YRRPYG 47  
 DB 444 YRRPYG 449

RESULT 37  
 SYE\_YERPE STANDARD; PRT; 471 AA.  
 AC Q8ZCK0;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)  
 GN (Gluts).  
 OS Yersinia pestis.  
 OS GLTX OR YPO2984.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 / Biovar Orientalis;  
 RX MEDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Fitzhugh R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebatina M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Fellwell T., Hamlin N., Holtroyd S., Jørgensen K., Kariyeh A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.,  
 RA "Genome sequence of Yersinia pestis, the causative agent of plague.",  
 RT Nature 413:523-527(2001).  
 RL Nature 413:523-527(2001).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +  
 CC di-phosphate + L-glutamyl-tRNA(Glu).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ414154; CAC92228.1; -  
 DR InterPro: IPR004527; GltX\_bact.  
 DR InterPro: IPR000924; GltX\_tRNA-synt\_1c.  
 DR InterPro: IPR001412; tRNA-synt\_1c.  
 DR Pfam: PF00749; tRNA-synt\_1c; 1.  
 DR PRINTS: PR00987; TRNASYNTHGLD.  
 DR TIGRFAms: TIGR00464; gltX\_bact; 1.  
 DR PROSITE: PS00178; AA\_tRNA\_LIGASE\_1; 1.  
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 9 19 "HIGH" REGION.  
 FT SITE 237 241 "KMSK" REGION.  
 FT BINDING 240 240 ATP (BY SIMILARITY).  
 SQ SEQUENCE 471 AA: 53110 MW: 6176293F008B358B CRC64;

Query Match 12.2%; Score 6; DB 1; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVP 11  
 DB 217 LGAPVP 222

RESULT 38  
 ICP0\_HSVB STANDARD; PRT; 532 AA.  
 AC P28980;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Trans-acting transcriptional protein ICP0.  
 GN 63.  
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicelloviruses.  
 OX NCBI\_TaxID=31520;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92295566; PubMed=1318606;  
 RA Teiford E.A.R., Watson M.S., McBride K., Davison A.J.;  
 RT "The DNA sequence of equine herpesvirus-1.";  
 RL Virology 189:304-316(1992).  
 RN [2]  
 RP STRUCTURE BY NMR OF 1-63.  
 RX MEDLINE=94087718; PubMed=8263911;  
 RA Everett R.D., Barlow P.N., Milner A., Luisi B., Orr A., Hope G.,  
 RA Lyon D.;  
 RT "A novel arrangement of zinc-binding residues and secondary structure  
 RT in the C3HC4 motif of an alpha herpes virus protein family.";  
 RL J. Mol. Biol. 234:1038-1047(1993).  
 RN [3]  
 RP STRUCTURE BY NMR OF 1-63.  
 RX MEDLINE=94172642; PubMed=8126734;  
 RA Barlow P.N., Luisi B., Milner A., Elliott M., Everett R.D.;  
 RT "Structure of the C3HC4 domain by 1H-nuclear magnetic resonance  
 RT spectroscopy. A new structural class of zinc-finger.";  
 RL J. Mol. Biol. 237:201-211(1994).  
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ICP0 PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M86664; AAB02498.1; -  
 DR PIR: I35801; WZBER5.  
 DR PDB: 1CHC; 3O-APR-94.  
 DR InterPro: IPR001841; ZnF\_Ring.  
 DR Pfam: PF00097; zf-C3HC4; 1.  
 DR SMART: SM00184; RING; 1.  
 DR PROSITE: PS00518; ZF\_RING\_1; 1.  
 DR PROSITE: PS00089; ZF\_RING\_2; 1.  
 KW Transcription regulation; Trans-acting factor; Activator; Zinc-finger;  
 KM DNA-binding; 3D-structure.  
 FT ZN\_FING 8 47 RING-TYPE.  
 FT METAL 8 8 ZINC 1.  
 FT METAL 11 11 ZINC 1.  
 FT METAL 24 24 ZINC 2.  
 FT METAL 26 26 ZINC 2.  
 FT METAL 29 29 ZINC 1.  
 FT METAL 32 32 ZINC 1.  
 FT METAL 43 43 ZINC 2.  
 FT METAL 46 46 ZINC 2.

FT DOMAIN 210 217 POLY-SER.  
 FT STRAND 19 21  
 FT TURN 22 25  
 FT STRAND 26 28  
 FT TURN 30 31  
 FT HELIX 32 39  
 FT TURN 43 46  
 FT STRAND 53 54  
 SQ SEQUENCE 532 AA; 58629 MW; BACB7E16FA26FDFA CRC64;

Query Match  
 Best Local Similarity 12.2%; Score 6; DB 1; Length 532;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 9 PVPYPD 14  
 |||||  
 Db 482 PVPYPD 487

RESULT 39  
 NOEB RHIME  
 ID NOEB RHIME STANDARD: PRT: 553 AA.  
 AC 052893:  
 DT 30-MAY-2000 (rel. 39, Created)  
 DT 30-MAY-2000 (rel. 39, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE Modulation protein noeb.  
 GN NOEB OR RA0417 OR SMA0774.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Plasmid pSYMA (megaplasmid 1).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RCR2011 / SU47;  
 RX MEDLINE=96111489; PubMed=8801423;  
 RA Ardourel M., Rortet G., Maillet F., Roche P., Truchet G.,  
 RA Prone J.-C., Rosenberg C.;  
 RA "In Rhizobium meliloti, the operon associated with the nod box n5  
 RT comprises nodI, noeA and noeb, three host-range genes specifically  
 RT required for the nodulation of particular Medicago species.";  
 RL Mol. Microbiol. 17:687-699(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396509; PubMed=11481432;  
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,  
 RA Barloy-Hubler F., Bowser L., Capela D., Gilbert F., Gouzy J.,  
 RA Gurjal M., Hong A., Hultzer L., Hyman R.W., Kahn D., Kahn M.L.,  
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
 RT "Nucleotide sequence and predicted functions of the entire  
 RT Sinorhizobium meliloti pSYMA megaplasmid.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).  
 CC -1- FUNCTION: NOT KNOWN: DOES NOT SEEM TO PARTICIPATE IN NOD FACTOR  
 CC SYNTHESIS BUT REQUIRED FOR NODULATION ON SOME SPECIFIC MEDICAGO  
 CC SPECIES SUCH AS M.LITTORALIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U26430; AAC44092.1; -  
 CC EMBL: AE007232; AAK65075.1; -  
 DR Plasmid: Nodulation; Transmembrane; Complete proteome.  
 KW TRANSMEM 4 24 POTENTIAL.  
 FT TRANSMEM 34 54 POTENTIAL.

FT TRANSMEM 65 85 POTENTIAL.  
 FT TRANSMEM 99 119 POTENTIAL.  
 FT TRANSMEM 122 142 POTENTIAL.  
 FT TRANSMEM 149 169 POTENTIAL.  
 SQ SEQUENCE 553 AA; 60732 MW; 1296ABD0222995F4 CRC64;

Query Match  
 Best Local Similarity 12.2%; Score 6; DB 1; Length 553;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 LADHTG 37  
 |||||  
 Db 88 LADHTG 93

RESULT 40  
 ECHP\_CAVPO  
 ID ECHP\_CAVPO STANDARD: PRT: 725 AA.  
 AC P55100;  
 DT 01-OCT-1996 (rel. 34, Created)  
 DT 01-OCT-1996 (rel. 34, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE Peroxisomal bifunctional enzyme (PBE) (PBEF) [includes: Enoyl-CoA  
 DE hydratase (EC 4.2.1.17); 3,2-trans-enoyl-CoA isomerase (EC 5.3.3.8);  
 DE 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)].  
 GN EHHADH.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=96140708; PubMed=8549802;  
 RA Cairn F., Cherkaoui-Malki M., Hoefler G., Latruffe N.;  
 RT "Cloning and tissue expression of two cDNAs encoding the peroxisomal  
 RT 2-enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase in the guinea  
 RT pig liver.";  
 RL FEBS Lett. 378:57-60(1996).  
 CC -1- CATALYTIC ACTIVITY: (3S)-3-hydroxyacyl-CoA - trans-2(or 3)-enoyl-  
 CC CoA + H(2)O.  
 CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxyacyl-CoA + NAD(+) = 3-oxoacyl-CoA  
 CC + NADH.  
 CC -1- CATALYTIC ACTIVITY: 3-cis-dodecenoyl-CoA = 2-trans-dodecenoyl-CoA.  
 CC -1- PATHWAY: CATALYZES THREE STEPS OF THE FATTY ACID BETA-OXIDATION  
 CC CYCLE.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ENOYL-COA  
 CC HYDRATASE/ISOMERASE FAMILY.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE 3-  
 CC HYDROXYACYL-COA DEHYDROGENASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X92742; CAA63403.1; -  
 DR EMBL: X85112; CAA59431.1; -  
 DR HSSP: P14604; 2DUB  
 DR InterPro: IPR002135; 3HCDH.  
 DR InterPro: IPR001753; EnCoA\_hydratase.  
 DR pfam: PF00378; ECH; 1.  
 DR pfam: PF00725; 3HCDH; 2.  
 DR pfam: PF02737; 3HCDH.N; 1.  
 DR PROSITE: PS00166; ENOYL-COA HYDRATASE; 1.  
 DR PROSITE: PS00342; MICROBODIES\_CTER; 1.  
 KW Fatty acid metabolism; Multifunctional enzyme; Oxidoreductase; NAD;  
 KW Lyase; Isomerase; Peroxisome.



FT	INIT_MET	0	0	BY SIMILARITY.
FT	DOMAIN	1	283	ENOTI-COA HYDRATASE / ISOMERASE.
FT	DOMAIN	284	574	3-HYDROXYACYL-COA DEHYDROGENASE.
FT	ACR_SITE	105	105	ACID (BY SIMILARITY).
FT	ACR_SITE	125	125	ACID (BY SIMILARITY).
FT	SITE	723	725	MICROBODY TARGETING SIGNAL (POTENTIAL).
SO	SEQUENCE	725 AA;	79243 MW;	645FC94CAF314E17 CRC64;
Qy	15 PLEPRR 20	12.2%;	Score 6;	DB 1;
Db	193 PLEPRR 198	Best Local Similarity 100.0%;	Pred. Nb. 50;	Matches 6;
		Conservative 0;	Mismatches 0;	Indels 0;
				Gaps 0;
RESULT 41				
ID	DPD4_BOVIN	STANDARD:	PRT:	765 AA.
AC	P81425; Q8WMC8;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DE	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Dipeptidyl peptidase IV (EC 3.4.14.5) (DPP IV) (T-cell activation antigen CD26) (Adenosine deaminase complexing protein) (ADCP-I)			
DE	(Activation molecule 3) (ACT3) (WC10).			
GN	DPD4 OR CD26.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID:9913;			
SN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymphocytes;			
RX	PubMed-12073152;			
RA	Lee S.-U., Park Y.-H., Davis W.C., Hamilton M.J., Naessens J.,			
RA	Bohach G.A.;			
RT	"Molecular characterization of bovine CD26 upregulated by a			
RT	staphylococcal superantigen.";			
RL	Immunogenetics 54:216-220(2002).			
RP	[2]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-17.			
RC	TISSUE=Thymus;			
RX	PubMed-11981836;			
RA	Gliddon D.R., Howard C.J.;			
RT	"CD26 is expressed on a restricted subpopulation of dendritic cells in			
RT	vitro.";			
RL	Eur. J. Immunol. 32:1472-1481(2002).			
RN	[3]			
RP	SEQUENCE OF 1-24.			
RC	TISSUE=T-cell;			
RX	PubMed-11598101;			
RA	Lee S.-U., Ferens W., Davis W.C., Hamilton M.J., Park Y.-H., Fox L.K.,			
RA	Naessens J., Bohach G.A.;			
RT	"Identity of activation molecule 3 on superantigen-stimulated bovine			
RT	cells is CD26.";			
RL	Infect. Immun. 69:7190-7193(2001).			
RN	[4]			
RP	SEQUENCE OF 537-546.			
RC	TISSUE=Kidney;			
RX	MEDLINE-98293306; PubMed-9629661;			
RA	Ben-Shoshan I., Parola A.H.;			
RT	"The CP-I subunit of adenosine deaminase complexing protein from calf			
RT	kidney is identical to human, mouse, and rat dipeptidyl peptidase			
RT	IV.";			
RL	Comp. Biochem. Physiol. 119B:289-292(1998).			
CC	-1- FUNCTION: Removes N-terminal dipeptides sequentially from			
CC	polypeptides having unsubstituted N-termini provided that the			
CC	penultimate residue is proline. Binds and regulates the activity			
CC	of ADA.			
CC	-1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Xbb- -			

```

OS      Human cytomegalovirus (strain AD169).
OC      Viruses; dsDNA viruses; no RNA stage; Herpesviridae;
OC      Betaherpesvirinae; Cytomegalovirus.
RN      NCBI_TaxID=10360;
[1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87169717; PubMed=3031311;
RA      Weston K., Barrell B.G.;
RT      "Sequence of the short unique region, short repeats, and part of the
RL      long repeats of human cytomegalovirus.";
RN      J. Mol. Biol. 192:177-208(1986).
[2]
RP      COMPLETE GENOME.
RX      MEDLINE=90269039; PubMed=2161319;
RA      Chae M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA      Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA      Predelli E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT      "Analysis of the protein-coding content of the sequence of human
RL      cytomegalovirus strain AD169.";
CC      Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC      - SIMILARITY: BELONGS TO THE US22 FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; X17403; CAN35269.1; -.
DR      EMBL; X04650; CAB37121.1; -.
DR      PIR; C27349; Q0BEE3.
DR      PIR; S09951; S09951.
DR      InterPro; IPR003360; US22.
DR      Pfam; PF02393; US22; 1.
DR      Hypothetical protein.
FT      CAROHD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CAROHD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CAROHD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE 788 AA; 83981 MW; 604B8C9C4472BC7A CRC64;

Query Match 12.2%; Score 6; DB 1; Length 788;
Matches 6; Conservative 0; Mismatches No. 54;

OY      10 VEYPDP 15
Db      758 VEYPDP 763

RESULT 43
PLSB_ECO57
ID      PLSB_ECO57 STANDARD; PRT; 806 AA.
AC      P58130;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Glycero1-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).
GN      PLSB OR Z5640 OR ECS5024.
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=83334;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX      MEDLINE=21074935; PubMed=11206551;
RA      Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA      Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoustis K.,
RA      Apodaca E.J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

```

```

RA  Weich R.A., Blattner F.R.;
RT  "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL  Nature 409:529-533(2001).
RN  (2)
RP  SEQUENCE FROM N.A.
RC  STRAIN=O157:H7 / RIMD 0509952;
RX  MEDLINE=0156231; PubMed=11258796;
RA  Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA  Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA  Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA  Kihara S., Shiba T., Hattori M., Shunagawa H.;
RT  "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT  O157:H7 and genomic comparison with a laboratory strain K-12.";
RL  DNA Res. 8:11-22(2001).
CC  -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC  acyl-sn-glycerol 3-phosphate.
CC  -1- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO
CC  FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC  -1- SUBCELLULAR LOCATION: Membrane-bound.
CC  -1- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation ~
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb.ch/announce/
CC  or send an email to license@isb.slb.ch).
CC  -----
DR  EMBL; AE005637; AAC59240.1; ALT_INIT.
DR  EMBL; AP002568; BAB38447.1; ALT_INIT.
DR  InterPro; IPR002123; Acyltransferase.
DR  Pfam; PF01553; Acyltransferase; 1.
KW  Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;
KW  Complete proteome.
FT  INT_MFR 0
FT  BY SIMILARITY.
SO  SEQUENCE 806 AA; 91278 MW; B07F39B876F5E5D6 CRC64;

Query Match. 12.2%; Score 6; DB 1; Length 806;
Best Local Similarity 100.0%; Pired. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 PDPLEP 18
Db 65 PDPLEP 70

RESULT 44
PLSB_ECOLI
ID PLSB_ECOLI STANDARD; PRT; 806 AA.
AC P00482; O9S683;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).
GN PLSB OR B4041.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=83291031; PubMed=6309817;
RA Lightner V.A., Bell R.M., Modrich P.;
RT "The DNA sequences encoding pIsb and dgk loci of Escherichia coli.";
RL J. Biol. Chem. 258:10856-10861(1983).
RN (2)
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the

```

```

RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
RN (3)
RP SEQUENCE FROM N.A. (MUTANT PLSB26).
RC STRAIN-BB26; PubMed=10074094;
RX MEDLINE=99173924; PubMed=10074094;
RA Heath R.J., Rock C.O.;
RT "A missense mutation accounts for the defect in the glycerol-3-
RT phosphate acyltransferase expressed in the plsb26 mutant.",
RL J. Bacteriol. 181:1944-1946(1999).
RN (4)
RP SEQUENCE OF 780-806 FROM N.A.
RC STRAIN-K12;
RX MEDLINE=92355505; PubMed=1644758;
RA Nichols B.P., Green J.M.;
RT "Cloning and sequencing of Escherichia coli uidC and purification of
RT chorismate lyase.",
RL J. Bacteriol. 174:5309-5316(1992).
RN (5)
RP PARTIAL SEQUENCE.
RX MEDLINE=83291032; PubMed=6350296;
RA Green P.R., Vannan T.C., Modrich P., Bell R.M.;
RT "Partial NH2- and COOH-terminal sequence and cyanogen bromide peptide
RT analysis of Escherichia coli sn-glycerol-3-phosphate
RT acyltransferase."
RL J. Biol. Chem. 258:10862-10866(1983).
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO
CC FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K00127; AAA24395.1; -
DR EMBL: U00006; AAC43135.1; ALT_INIT.
DR EMBL: AE000477; AAC77011.1; ALT_INIT.
DR EMBL: AF106625; AAD20588.1; -
DR EMBL: M93413; AAA24718.1; -
DR EMBL: M93136; AAA24713.1; -
DR PIR: A00565; XUECAG.
DR PIR: C42956; C42956.
DR ECGene: EG10740; PLSB.
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;
KW Complete proteome.
FT INIT_MET 0
FT VARIANT 348 348 A -> T (IN PLSB26; RESULTS IN HIGH KM FOR
FT GLYCEROL-3-PHOSPHATE AND REDUCED SPECIFIC
FT ACTIVITY).
SQ SEQUENCE 806 AA; 91250 MW; 407F2EB878D16EBB CRC64;

```

```

Query Match 12.2%; Score 6; DB 1; Length 806;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 13 PDPLEP 18
DB 65 PDPLEP 70

```

```

RESULT 45
PLSB_SALTI STANDARD; PRT; 806 AA.
ID PLSB_SALTI
AC 082176;

```

```

DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).
GN PLSB OR STY431.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBL_TaxID=601;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Brooks R.M., Dowd L., Farrar J.,
RA Cronin A., Davis P., Davies R.M., Dord L., White N., Farrar J.,
RA Feltham T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar typhi CT18."
RL Nature 413:848-852(2001).
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO
CC FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL627282; CAD09219.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;
KW Complete proteome.
SQ SEQUENCE 806 AA; 91241 MW; DD950CD4D80CEB880 CRC64;

```

```

Query Match 12.2%; Score 6; DB 1; Length 806;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 13 PDPLEP 18
DB 66 PDPLEP 71

```

```

Search completed: December 4, 2002, 15:39:00
Job time : 13 secs

```



## RESULT 2

ID	Q99K39	PRELIMINARY;	PRT;	95 AA.
AC	Q99K39			
DT	01-JUN-2001 (TREMBLrel, 17, Created)			
DT	01-JUN-2001 (TREMBLrel, 17, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel, 21, Last annotation update)			
DE	Similar to bone gamma-carboxylutamate protein, related sequence 1.			
GN	A1461847.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Strausberg R.;			
RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; BC005483; AAH05483.1; -.			
DR	MGI; MGI:2139729; A1461847.			
DR	InterPro: IPR002384; GLA_bone.			
DR	InterPro: IPR000294; VitK_dep_GLA.			
DR	Pfam; PF00594; gla; 1.			
DR	PRINTS; PR00002; GLABONE.			
DR	SMART; SMO0069; GLA; 1.			
DR	PROSITE; PS00011; GLU CARBOXYLATION; 1.			
SO	SEQUENCE 95 AA; 10445 MW; 25C4AB3AA8B51909E CRC64;			

	Query Match	Similarity	Score	DB	Length
Best Local	56.5%	Pred. No. 3e-12			95
Matches	26	Conservative	7	Mismatches	13
				Indels	0
				Gaps	0
QY	2	LYQMGAPVPPDPLXPRRXVCLMDDELADHDFGOFARREYG	47		
DB	47	LRRIYGVSPDPLPLEPTRELCEDLPACCELSNÖGKTAIRRTG	92		

RESULT 3			
ID	Q90YW2	PRELIMINARY:	PRT: 97 AA.
AC	Q90YW2		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Osteocalcin.		
OS	Sparus aurata (gilthead sea bream).		
OC	Eukaryota; Metazoa; Chordata; Cratiata; Vertebrata; Euteleostomi;		
OC	Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
OC	Acanthomorphi; Acanthopterygii; Percomorpha; Perciformes; Percoidae;		
OC	Sparidae; Sparus.		
OX	NCBI_TaxID=8175;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21297182; PubMed=11404005;		
RA	Pinto J.P., Ohresser M.C.P., Cancela M.L.;		
RT	"Cloning of the bone Gla protein gene from the teleost fish Sparus		
RT	aurata. Evidence for overall conservation in gene organization and		
RT	bone-specific expression from fish to man.";		
RL	gene 270:77-91(2001).		
DR	EMBL, AF048703; AAK6568.1; -.		
DR	EMBL, AF289506; AAK62679.1; -.		
DR	InterPro: IPR000294; VitK_dep_GLA.		
DR	Pfam: PF00594; gla; 1.		
DR	PROSITE, PS00011; GLU_CARBOXYLATION; UNKNOWN.1.		
SQ	SEQUENCE 97 AA: 10434 MW: B95608824FPFEECB CRC64;		

Query Match	26.1%	Score 71:	DB 13:	Length 97:
Best Local Similarity	43.3%	Pred. No.	0.035:	
Matches 13:	Conservative	4:	Mismatches 13:	Indels 0:
			Gaps	0

QY 20 RYXCLNPDCDELADHIGFQEAIRRFYGPV 43  
| | | | : : | | : : : :  
DB 66 REVCELNLACEHMDTEGIIAAYTAYYGPV 95

RESULT 4		
ID	092553	PRELIMINARY; PRJ: 147 AA.
AC	Q92553	
DT	01-DEC-2001 (TREMBlrel. 19, Created)	
DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)	
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)	
DE	MRP3.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-BALB/C;	
RX	PubMed=11819679;	
RA	Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,	
RA	Su C.;	
RT	"Mechanism of exogenous nucleic acids and their precursors improving	
RT	the repair of intestinal epithelium after irradiation in mice.";	
RT	World J. Gastroenterol. 6:709-717(2000).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-BALB/C;	
RA	Cui D., Zeng G., Yan X., Li X., Su C.;	
RT	"Cloning of mouse genes related to repairing of intestinal epithelium	
RT	of the irradiated mice by treatment with the intestinal RNA of mice of	
RT	the same strain.";	
RT	Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).	
DR	EMBL; AF240166; AAK43731.1; .	
DR	InterPro; IPR003006; IG_MHC.	
DR	Pfam; PF00047; 19; 1.	
SO	SEQUENCE 147 AA; 16274 MW; 800594A12B97191F CRC64;	

Query Match	25.7%	Score 70:	DB 11:	Length 147:
Best Local Similarity	92.3%	Pred. No.	0.075:	
Matches	12:	Conservative	0:	Mismatches 1:
				Indels 0:
				Gaps 0:
QY	7	GAPVPPPLXPR	19	
DB	133	GAPVPPPLXPR	145	

RESULT 5		
092552		
ID	Q92552	PRELIMINARY;
AC	Q92552	PRT; 170 AA.
DT	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)	
DE	MRP4.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
OX	NCBI_TaxId=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BALB/C;	
RX	PubMed=11819679;	
RA	Sui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,	
RA	Su C.;	
RT	"Mechanism of exogenous nucleic acids and their precursors improving	
RT	the repair of intestinal epithelium after irradiation in mice.";	
RL	World J. Gastroenterol. 6:709-717(2000).	

RN [2]  
RP SEQUENCE FROM N.A.  
RQ STRAIN-BALB/C;  
RA Cui D., Zeng G., Yan X., Li X., Su C.;  
RB "Cloning of mouse genes related to repairing of intestinal epithelium  
RT of the irradiated mice by treatment with the intestinal RNA of  
Rt the same strain";  
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).



OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Apodaca E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoculis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
DR EMBL: AE005437; AAG57175.1; -;  
DR EMBL: AP002560; BAB36343.1; -;  
DR InterPro: IPR004495; MetG\_Cterm.  
DR InterPro: IPR002300; tRNA-synt\_1a.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002304; tRNA-synt\_met.  
DR InterPro: IPR002547; tRNA\_bind.  
DR Pfam: PF00133; tRNA-synt\_1; 1.  
DR Pfam: PF01588; tRNA\_bind; 1.  
DR PRINTS: PRO1041; TRNASYNTHET.  
DR TIGRFAMS: TIGR00398; metG; 1.  
DR TIGRFAMS: TIGR00399; metG\_C-term; 1.  
DR PROSITE: PS00178; AA:TRNA\_LIGASE\_I; 1.  
KW Aminoacyl-tRNA synthetase; Complete proteome.  
SQ SEQUENCE 677 AA; 76254 MW; DBAF4C76AE04FCA CRC64;  
  
Query Match 21.9%; Score 59.5; DB 16; Length 677;  
Best Local Similarity 27.3%; Pred. No. 11;  
Matches 12; Conservative 7; Mismatches 16; Indels 9; Gaps 1;  
  
OY 1 YIYQNLGAPVPPDPLPRXYCXNLPCDDELADHIGFOEAYR 44  
DB 250 YFYWLDAPIGY-----MGSPKLCIDKRGDSVSPDEYWK 284  
  
RESULT 10  
O18313  
ID O18313 PRELIMINARY; PRT; 706 AA.  
AC O18313;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Hypothetical 82.6 kDa protein.  
GN C29H12.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Favello A.;  
RT "The sequence of C. elegans cosmid C29H12.";  
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 'CHROMO' DOMAIN.  
DR EMBL: U23169; AAC46597.1; -;  
DR InterPro: IPR00953; Chromo.  
DR InterPro: IPR00719; Euk\_pkinase.  
DR Pfam: PF00385; Chromo; 1.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00298; CHROMO; 1.  
DR PROSITE: PS00598; CHROMO\_1; 1.  
DR PROSITE: PS50013; CHROMO\_2; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW ATP-binding; Hypothetical protein; Nuclear protein; Transferase.  
SQ SEQUENCE 706 AA; 82646 MW; 74E3EF3E109EAFCC CRC64;  
  
Query Match 21.7%; Score 59; DB 5; Length 706;  
Best Local Similarity 45.0%; Pred. No. 13;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
  
OY 28 DCDDELADHIGFOEAYRFFYG 47  
DB 47 EIERIDHVSFLAEAVESFYG 66  
  
RESULT 11  
O43413  
ID O43413 PRELIMINARY; PRT; 369 AA.  
AC O43413;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypothetical 41.8 kDa protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=96207227; PubMed=8619474;  
RA Andersson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;  
RT "A 'double adaptor' method for improved shotgun library  
RT construction.";  
RL Anal. Biochem. 236:107-113(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=97264341; PubMed=9110174;  
RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,  
RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;  
RT "Large-scale concatenation cDNA sequencing.";  
RL Genome Res. 7:353-358(1997).  
DR EMBL: AF007156; AAC19157.1; -;  
DR HSSP: P21707; IBYN.  
DR InterPro: IPR000008; C2.  
DR InterPro: IPR001478; PDZ.  
DR InterPro: IPR001565; Synaptotagmin.  
DR Pfam: PF00168; C2; 1.  
DR Pfam: PF00595; PDZ; 1.  
DR PRINTS: PR00399; SYNAPTOTAGMN.  
DR SMART: SM00239; C2; 1.  
DR SMART; SM00228; PDZ; 1.



DR PROSITE: PS50004; C2\_DOMAIN\_2; 1.  
 DR PROSITE: PS50106; PDZ; 1.  
 KM Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 369 AA; 41822 MW; 613PFB065C7FAA8C CRC64;

Query Match 21.5%; Score 58.5; DB 4; Length 369;  
 Best Local Similarity 30.2%; Pred. No. 7.8;  
 Matches 16; Conservative 4; Mismatches 12; Indels 21; Gaps 2;

OY 9 PVPYPPPLXRRVYCXKLN-----DCDELADHIGFQEAAYR 43  
 DB 270 PLPHPSPYMRRLQHGESPTRLQSRKISDSVSDYDCD---DGIQVSDYR 319

## RESULT 12

ID 09U026 PRELIMINARY; PRT; 1188 AA.

AC 09U026;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE KIA0751 protein.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;

RN (1)  
 RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;  
 RX MEDLINE=99087487; PubMed=9872452;

RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,  
 Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XI.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";

RL DNA Res. 5:277-286(1998).  
 DR EMBL: AB018294; BAA34471.1; -.

DR HSSP: P21707; IBYI.

DR InterPro: IPR000008; C2.

DR InterPro: IPR001478; PDZ.

DR Pfam: PF00168; C2; 2.

DR Pfam: PF00595; PDZ; 1.

DR SMART: SM00239; C2; 2.

DR SMART: SM00228; PDZ; 1.

DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.

DR PROSITE: PS50106; PDZ; 1.

SQ SEQUENCE 1188 AA; 135250 MW; DD65DB3228055AF CRC64;

Query Match 21.5%; Score 58.5; DB 4; Length 1188;  
 Best Local Similarity 30.2%; Pred. No. 27;  
 Matches 16; Conservative 4; Mismatches 12; Indels 21; Gaps 2;

OY 9 PVPYPPPLXRRVYCXKLN-----DCDELADHIGFQEAAYR 43  
 DB 716 PLPHPSPYMRRLQHGESPTRLQSRKISDSVSDYDCD---DGIQVSDYR 765

## RESULT 13

ID 09IUM9 PRELIMINARY; PRT; 106 AA.

AC 09IUM9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Trm protein.

DE Trm protein.

GN TmM.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Plasmid PSB102.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI\_TaxID=382;

RN (1)  
 RP SEQUENCE FROM N.A.

RC TRANSPOSON-MERCURY RESISTANCE TRANSPOSON TN5718;

RA Schneider S., Keller M., Droege M., Lanka E., Puhler A.,

RA Selbitschka W.;

RT "The genetic organization and evolution of the broad-host-range

RT mercury resistance plasmid PSB102 isolated from a microbial population

RT residing in the rhizosphere of alfalfa.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ304453; CAC79208.1; -.

DR InterPro: IPR001633; EAL.

DR Pfam: PF00563; EAL; 1.

KW Plasmid.

SQ SEQUENCE 106 AA; 11420 MW; D744FF67C546E214 CRC64;

Query Match 21.3%; Score 58; DB 2; Length 106;  
 Best Local Similarity 28.3%; Pred. No. 2.4;  
 Matches 15; Conservative 6; Mismatches 24; Indels 8; Gaps 1;

OY 1 YLVOWIGAPVPPDPLXPRXVC-----XINPDCDELADHIGFQEAAYRF 45  
 DB 31 FVSQWKGATMANADPRTATSCVCCCKEIPDAATPBGAEYVHFCCGYORF 83

## RESULT 14

ID 08VHV2 PRELIMINARY; PRT; 196 AA.

AC 08VHV2;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Noda protein.

GN NODA.

OS Mesorhizobium plurifarium.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI\_TaxID=69974;

RN (1)  
 RP SEQUENCE FROM N.A.

RC STRAIN=ORS1096;

RA Ba S., Willems A., Lorquin J., Roche P., de Lajudie P., Neyra M.,  
 Moulin L., Gillis M., Dreyfus B., Boivin-Masson C.;

RT "Symiotic and taxonomic diversity of Acacia tortilis rhizobia.";

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ302678; CAC80541.1; -.

DR InterPro: IPR003484; NODA.

DR Pfam: PF02474; NODA; 1.

DR PROSITE: PS01349; NODA; 1.

SQ SEQUENCE 196 AA; 21708 MW; 9BAA7032FAB1EEC0 CRC64;

Query Match 21.3%; Score 58; DB 2; Length 196;  
 Best Local Similarity 55.6%; Pred. No. 4.6;  
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 31 ELADHIGFQEAAYRFYGP 48  
 DB 16 ELSDHVELAEFFRTYGP 33

## RESULT 15

ID 09FW3 PRELIMINARY; PRT; 113 AA.

AC 09FW3;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE Urf2 protein.

DE Urf2 protein.

GN URF2.

OS Pseudomonas sp. ED23-33.

OC Bacteria; Proteobacteria.

OX NCBI\_TaxID=83781;

RN (1)  
 RP SEQUENCE FROM N.A.

RC TRANSPOSON-TN5058;  
RA Minakhina S., Minakhin L., Kholodil G., Mindlin S., Gorlenko Z.H.,  
RA Yurleva O., Nikiforov V.;  
RT "Molecular inventory of transposons from environmental bacteria:  
RT epidemic dissemination of Tn21-, Tn5041-, and Tn5053-related  
RT transposons.";  
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Y17897; CAC14707.1; -;  
SQ SEQUENCE 113 AA; 11996 MW; 17A059D39BB2400E CRC64;

Query Match 21.1%; Score 57.5; DB 2; Length 113;  
Best Local Similarity 27.8%; Pred. No. 3;  
Matches 15; Conservative 6; Mismatches 24; Indels 9; Gaps 1;

Oy 1 YLYQWLGAPVPPDPLXPRXYC-----XLPDDELADHIGFOEAYRRF 45  
::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
Db 37 FVQNRGATMNDPTATSCVQWSKEIPLDAFTPEGAEVVEHFCGEYORF 90

Search completed: December 4, 2002, 15:45:35  
Job time : 32 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:40:30 ; Search time 11 Seconds

(without alignments)  
184.758 Million cell updates/sec

Title: US-09-462-931-2-COPY

Perfect score: 272  
Sequence: 1 YIXQNLGAPVPPDPLXPRR.....DELDHIGFQEAIRRYGCV 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266	97.8	100	1	OSTC_HUMAN
2	259	95.2	100	1	P02818 homo sapien
3	247	90.8	49	1	OSTC_MACFA
4	244	89.7	100	1	OSTC_HORSE
5	226	83.1	49	1	OSTC_BOVIN
6	219	80.5	49	1	OSTC_CANFA
7	202	74.3	49	1	OSTC_FELCA
8	199	73.2	49	1	OSTC_RAT
9	181	66.5	48	1	OSTC_DRONO
10	168.5	61.9	97	1	OSTC_CHICK
11	168	61.8	49	1	OSTC_XENLA
12	148	54.4	95	1	OSTC_MOUSE
13	144	52.9	95	1	OSTC_MOUSE
14	91	33.5	47	1	OSTC_XIPGL
15	71	26.1	45	1	OSTC_SPAAU
16	68	25.0	45	1	OSTC_LEPMA
17	60	22.1	344	1	GUNA_TREPE
18	59.5	21.5	676	1	SYM_ECOLI
19	58.5	21.5	203	1	PAXI_CHICK
20	57.5	21.1	423	1	TF3A_HUMAN
21	57.5	21.1	430	1	Y588 MENJA
22	57	21.0	393	1	ACK2_RHIME
23	56.5	20.8	435	1	DCOR_PANRE
24	56	20.6	262	1	RS4_GOSHI
25	55	20.6	264	1	RS4_SOLTU
26	55	20.2	102	1	MGP_GALGA
27	55	20.2	384	1	SYXB_HUMAN
28	54.5	20.0	139	1	YXCE_BACSU
29	54	19.9	111	1	FREE_HUMAN
30	54	19.9	210	1	MODA_BRASN
31	54	19.9	261	1	RS4_PRUAR
32	54	19.9	265	1	RS4_MAIZE
33	54	19.9	759	1	Z287_MOUSE

34	54	19.9	2326	1	PGG2_RAT
35	53.5	19.7	486	1	ENV_HTV2
36	53	19.5	291	1	USO2_HSV2H
37	53	19.5	726	1	RNR_MYCPN
38	52	19.1	111	1	FREE_MOUSE
39	52	19.1	352	1	SMP_SERMA
40	52	19.1	390	1	KC11_RAT
41	52	19.1	422	1	KC11_HUMAN
42	51	18.8	196	1	MODA_RHISN
43	51	18.8	210	1	MODA_BRASN
44	51	18.8	503	1	CP3P_MOUSE
45	50.5	18.6	361	1	PAXI_HUMAN

## ALIGNMENTS

## RESULT 1

ID	OSTC_HUMAN	STANDARD:	PRT:	100 AA.
AE	P02818;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)			
DE	(Bone Gla-protein) (BGP).			
GN	BGLAP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-90245603; PubMed-2336375;			
RA	Kiefer M.C., Saphire A.C.S., Bauer D.M., Barr P.J.;			
RT	"The cDNA and derived amino acid sequences of human and bovine bone			
RT	Gla protein.";			
RL	Nucleic Acids Res. 18:1909-1909(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-87004555; PubMed-3019668;			
RA	Celette A.J., Buecker J.L., Kriz R., Wang E.A., Wozney J.M.;			
RT	"Isolation of the human gene for bone gla protein utilizing mouse and			
RT	rat cDNA clones.";			
RL	EMBO J. 5:1885-1890(1986).			
RN	[3]			
RP	SEQUENCE OF 52-100.			
RX	MEDLINE-81006914; PubMed-6967872;			
RA	Poser J.W., Esch F.S., Ling N.C., Price P.A.;			
RT	"Isolation and sequence of the vitamin K-dependent protein from human			
RT	bone. Undercarboxylation of the first glutamic acid residue.";			
RL	J. Biol. Chem. 255:8685-8691(1980).			
CC	- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS			
CC	STRONGLY TO APATITE AND CALCIUM.			
CC	- P-TM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K			
CC	DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE			
CC	BINDING OF CALCIUM.			
CC	- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN			
CC	FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL: X53698; CAA37736.1; -			
CC	EMBL: X51699; CAA35996.1; -			
CC	EMBL: X04143; CAA27763.1; -			
CC	PIR: A03301; GEHU.			
CC	PIR: C25471; C25471.			

DR PIR: S12652; S12652.  
 DR Genew: HGNC:1043; BGLAP.  
 DR MIM: 112260; .  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA; 1.  
 DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.  
 DR Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone; Signal.  
 FT SIGNAL 1 23 PROBABLE.  
 FT PROPEP 24 51 PROBABLE.  
 FT CHAIN 52 100 OSTEOCALCIN.  
 FT MOD\_RES 68 68 GAMMA-CARBOXYGLUTAMIC ACID (IN 9% OF THE  
 FT MOD\_RES 72 72 MOLECULES).  
 FT MOD\_RES 75 75 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 74 80 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT CONFLICT 33 34 MISSING (IN REF. 2).  
 SQ SEQUENCE 100 AA; 10962 MW; 4DF2A0A0849CB71 CRC64;

Query Match 97.8%; Score 266; DB 1; Length 100;  
 Best Local Similarity 93.9%; Pred. No. 3.3e-29;  
 Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLPRRYXCXLPDCDELADHIGFQEAAYRRFGPV 49  
 |||||  
 DB 52 YLYQWLGAPVPPDPLPRRYXCXLPDCDELADHIGFQEAAYRRFGPV 100

RESULT 2  
 ID OSTC\_MACFA STANDARD; PRT: 49 AA.  
 AC P02819;  
 DT 21-JUL-1986 (Rel. 01; Created)  
 DT 21-JUL-1986 (Rel. 01; Last sequence update)  
 DT 15-JUN-2002 (Rel. 41; Last annotation update)  
 DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-  
 DE protein) (BGP).  
 GN BGLAP.  
 OS Maccaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=82182842; PubMed=6978733;  
 RA Hauschka P.V.; Carr S.A.; Blemann K.;  
 RT "Primary structure of monkey osteocalcin.";   
 RL Biochemistry 21:638-642(1982).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTH: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 CC PIR: A03302; GEMKI.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA; 1.  
 DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.  
 DR Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;  
 KW Bone.  
 FT MOD\_RES 9 9 HYDROXYLATION.  
 FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 23 29 BY SIMILARITY.  
 SQ SEQUENCE 49 AA; 5743 MW; C20116014D0C4958 CRC64;

Query Match 95.2%; Score 259; DB 1; Length 49;  
 Best Local Similarity 89.8%; Pred. No. 1.3e-28;  
 Matches 44; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLPRRYXCXLPDCDELADHIGFQEAAYRRFGPV 49  
 |||||  
 DB 1 YLYQWLGAPVPPDPLPRRYXCXLPDCDELADHIGFQEAAYRRFGPV 49

RESULT 3  
 ID OSTC\_HORSE STANDARD; PRT: 49 AA.  
 AC P83005;  
 DT 15-JUN-2002 (Rel. 41; Created)  
 DT 15-JUN-2002 (Rel. 41; Last sequence update)  
 DT 15-JUN-2002 (Rel. 41; Last annotation update)  
 DE Osteocalcin (gamma-carboxyglutamic acid-containing protein) (Bone GLA-  
 DE protein) (BGP).  
 GN BGLAP.  
 OS Equus caballus (horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE, HYDROXYLATION, AND GAMMA-CARBOXYGLUTAMIC ACIDS.  
 RC TISSUE=Bone;  
 RA Carstensen B., Mattiez R., Amory H., Lepage O.M., Remy B.;  
 RT "Isolation and characterization of equine osteocalcin.";   
 RL Ann. Med. Vet. 0:0-0(2002).  
 CC -1- FUNCTION: Constitutes 1-2% of the total bone protein. It binds  
 CC strongly to apatite and calcium.  
 CC -1- PTH: Gamma-carboxyglutamic acid residues are formed by vitamin K  
 CC dependent carboxylation. These residues are essential for the  
 CC binding of calcium.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR SMART: SM00069; GLA; 1.  
 DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.  
 KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K;  
 KW Hydroxylation.  
 FT MOD\_RES 9 9 HYDROXYLATION.  
 FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 23 29 BY SIMILARITY.  
 SQ SEQUENCE 49 AA; 5732 MW; A5B826014D12857F CRC64;

Query Match 90.8%; Score 247; DB 1; Length 49;  
 Best Local Similarity 87.8%; Pred. No. 5.6e-27;  
 Matches 43; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLPRRYXCXLPDCDELADHIGFQEAAYRRFGPV 49  
 |||||  
 DB 1 YLYQWLGAPVPPDPLPRRYXCXLPDCDELADHIGFQEAAYRRFGPV 49

RESULT 4  
 ID OSTC\_BOVIN STANDARD; PRT: 100 AA.  
 AC P02820;  
 DT 21-JUL-1986 (Rel. 01; Created)  
 DT 01-AUG-1990 (Rel. 15; Last sequence update)  
 DT 15-JUN-2002 (Rel. 41; Last annotation update)  
 DE Osteocalcin precursor (gamma-carboxyglutamic acid-containing protein)  
 DE (Bone Gla-protein) (BGP).  
 GN BGLAP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.

```

OX  NCBI_TaxID-9913;
RN  [1]
RX  MEDLINE-90245603; PubMed-2336375;
RA  Kiefer M.C., Sapphire A.C.S., Bauer D.M., Barr P.J.;
RT  "The cDNA and derived amino acid sequences of human and bovine bone
RL  Gla protein.";
RN  Nucleic Acids Res. 18:1909-1909(1990).
RN  (2)
RX  MEDLINE-7036749; PubMed-1068450;
RA  Price P.A., Poser J.W., Raman N.;
RT  "Primary structure of the gamma-carboxyglutamic acid-containing
RT  protein from bovine bone.";
RL  Proc. Natl. Acad. Sci. U.S.A. 73:3374-3375(1976).
CC  -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC  STRONGLY TO APATITE AND CALCIUM.
CC  -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC  DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC  BINDING OF CALCIUM.
CC  -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC  FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: X53699; CAA37737.1; -;
DR  EMBL: X51700; CAA35997.1; -;
DR  PIR: A03303; GEB0.
DR  PIR: S12653; S12653.
DR  InterPro: IPR002384; GLA_bone.
DR  InterPro: IPR000294; VitK_dep_GLA.
DR  Pfam: PF00594; gla; 1.
DR  PRINTS; PRO0002; GLABONE.
DR  SMART; SM00069; GLA; 1.
DR  PROSITE; PS00011; GLU-CARBOXYLATION; 1.
DR  Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;
KW  Bone; Signal.
FT  SIGNAL 1 23
FT  PROPEP 24 31 PROBABLE.
FT  CHAIN 52 100 OSTEOCALCIN.
FT  MOD_RES 60 60 HYDROXYLATION.
FT  MOD_RES 68 68 GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES 72 72 GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES 75 75 GAMMA-CARBOXYGLUTAMIC ACID.
FT  DISULFID 74 80
SQ  SEQUENCE 100 AA; 11042 MW; 73015D1681B26219 CRC64;

Query Match 89.7%; Score 244; DB 1; Length 100;
Best Local Similarity 85.7%; Pred. No. 3.2e-26;
Matches 42; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLXPRRXVXCLNPDDELADHIGFOEAYRFGVPV 49
   || ||||| ||||| : || ||||| ||||| ||||| |||||
DB 52 YLDHMLGAPARPPDPLEPKREVCELNPDCDELADHIGFOEAYRFGVPV 100

RESULT 5
OSTC_CANFA
ID OSTC_CANFA STANDARD; PRT; 49 AA.
AC P81455;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
GN BGLAP.
OS Canis familiaris (Dog).

```

```

OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX  NCBI_TaxID-9615;
RN  [1]
RX  MEDLINE-93318657; PubMed-8101026;
RA  Colombo G., Fantl P., Yao C., Malluche H.H.;
RT  "Isolation and complete amino acid sequence of osteocalcin from
RT  canine bone.";
RL  J. Bone Miner. Res. 8:733-743(1993).
CC  -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC  STRONGLY TO APATITE AND CALCIUM.
CC  -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC  DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC  BINDING OF CALCIUM.
CC  -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC  FAMILY.
DR  InterPro: IPR002384; GLA_bone.
DR  InterPro: IPR000294; VitK_dep_GLA.
DR  Pfam: PF00594; gla; 1.
DR  PRINTS; PRO0002; GLABONE.
DR  SMART; SM00069; GLA; 1.
DR  PROSITE; PS00011; GLU-CARBOXYLATION; 1.
KW  Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;
KW  Bone.
FT  MOD_RES 9 9 HYDROXYLATION (BY SIMILARITY).
FT  MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.
FT  DISULFID 23 29
SQ  SEQUENCE 49 AA; 5524 MW; 43121D015817CEA6 CRC64;

Query Match 83.1%; Score 226; DB 1; Length 49;
Best Local Similarity 81.6%; Pred. No. 3.9e-24;
Matches 40; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLXPRRXVXCLNPDDELADHIGFOEAYRFGVPV 49
   || ||||| ||||| : || ||||| ||||| ||||| |||||
DB 1 YLDHMLGAPVPPDPLEPKREVCELNPDCDELADHIGFOEAYRFGVPV 49

RESULT 6
OSTC_FELCA
ID OSTC_FELCA STANDARD; PRT; 49 AA.
AC P02821;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE protein) (BGP).
GN BGLAP.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID-9685;
RN [1]
RX MEDLINE-85054706; PubMed-6334077;
RA Shimomura H., Kanai Y., Sanada K.;
RT "Primary structure of cat osteocalcin.";
RL J. Biochem. 96:405-411(1984).
CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS
CC STRONGLY TO APATITE AND CALCIUM.
CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC BINDING OF CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC FAMILY.
DR PIR: A03304; GECT.
DR InterPro: IPR002384; GLA_bone.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00594; gla; 1.
DR PRINTS; PRO0002; GLABONE.

```

DR SMART; SM00069; GLA; 1.  
 DR PROSITE; PS00011; GLU-CARBOXYLATION; 1.  
 KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;  
 KW Bone  
 FT MOD\_RES 9 9 HYDROXYLATION.  
 FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 23 29  
 SQ SEQUENCE 49 AA; 5495 MW; 93D2131FA9F656D3 CRC64;  
 Query Match 80.5%; Score 219; DB 1; Length 49;  
 Best Local Similarity 77.6%; Pred. No. 3.5e-23;  
 Matches 38; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 OY 1 YLYQWLGAPVPYPPDLXRRXYCXLPNCDELADHIGQEAVERFYGV 49  
 DB 1 YLYQWLGAPVPYPPDLXRRXYCXLPNCDELADHIGQEAVERFYGV 49  
 RESULT 7  
 ID OSTC\_RAT STANDARD; PRT: 99 AA.  
 AC P04640;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)  
 DE (Bone Gla-protein) (BGP).  
 GN BGLAP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87004555; PubMed=3019668;  
 RA Celeste A.J., Buecker J.L., Kriz R., Wang E.A., Wozney J.M.;  
 RT "Isolation of the human gene for bone gla protein utilizing mouse and  
 RT rat cDNA clones.";  
 RL EMBO J. 5:1885-1890(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85298305; PubMed=3875856;  
 RA Pan L.C., Price P.A.;  
 RT "The propeptide of rat bone gamma-carboxyglutamic acid protein shares  
 RT homology with other vitamin K-dependent protein precursors.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6109-6113(1985).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89118266; PubMed=3265336;  
 RA Yoon K., Rutledge S.J.C., Buenaga R.F., Rodan G.A.;  
 RT "Characterization of the rat osteocalcin gene: stimulation of  
 RT promoter activity by 1,25-dihydroxyvitamin D3.";  
 RL Biochemistry 27:8521-8526(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89251082; PubMed=2785907;  
 RA Theofan G., Haberstroh L.M., Price P.A.;  
 RT "Molecular structure of the rat bone Gla protein gene and  
 RT identification of putative regulatory elements.";  
 RL DNA 8:213-221(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89145200; PubMed=2784002;  
 RA Lian J., Stewart C., Puchacz E., Mackowiak S., Shalhoub V.,  
 RA Collart D., Zambetti G., Stein G.;  
 RT "Structure of the rat osteocalcin gene and regulation of vitamin D-  
 RT dependent expression.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:1143-1147(1989).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K

CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X04141; CA27761.1; -;  
 DR EMBL; M11777; AAA40816.1; -;  
 DR EMBL; M23637; AAA41761.1; -;  
 DR EMBL; M25490; AAA53280.1; -;  
 DR EMBL; J04500; AAA41764.1; -;  
 DR PIR; A25167; GERT.  
 DR PIR; A25471; A25471.  
 DR PIR; A31856; A31856.  
 DR PIR; A31419; A31419.  
 DR PIR; A32324; A32324.  
 DR InterPro; IPR002384; GLA\_bone.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00594; gla; 1.  
 DR PRINTS; PRO0002; GLABONE.  
 DR SMART; SM00069; GLA; 1.  
 DR PROSITE; PS00011; GLU-CARBOXYLATION; 1.  
 KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;  
 KW Bone; Signal.  
 FT SIGNAL 1 23 PROBABLE.  
 FT PROPEP 24 49 PROBABLE.  
 FT CHAIN 50 99 OSTEOCALCIN.  
 FT MOD\_RES 58 58 HYDROXYLATION.  
 FT MOD\_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID  
 FT MOD\_RES 70 70 (BY SIMILARITY).  
 FT MOD\_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID  
 FT MOD\_RES 73 73 (BY SIMILARITY).  
 FT MOD\_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID  
 FT MOD\_RES 73 73 (BY SIMILARITY).  
 FT DISULFID 72 78 BY SIMILARITY.  
 FT SEQUENCE 99 AA; 10927 MW; 7F18F1866D4E4388 CRC64;  
 Query Match 74.3%; Score 202; DB 1; Length 99;  
 Best Local Similarity 74.5%; Pred. No. 1.5e-20;  
 Matches 35; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
 OY 1 YLYQWLGAPVPYPPDLXRRXYCXLPNCDELADHIGQEAVERFYGV 47  
 DB 50 YLYQWLGAPVPYPPDLXRRXYCXLPNCDELADHIGQEAVERFYGV 96  
 RESULT 8  
 ID OSTC\_RABIT STANDARD; PRT: 49 AA.  
 AC P39056;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein)  
 DE (Bone Gla-protein) (BGP).  
 GN BGLAP.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_Taxid=9986;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92175242; PubMed=1794506;  
 RA Viridi A.S., Willis A.C., Hauschka P.V., Triffitt J.T.;  
 RT "Primary amino acid sequence of rabbit osteocalcin.";  
 RL Biochem. Soc. Trans. 19:373S-373S(1991).

CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM  
 CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 DR PIR: A61280; A61280.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00594; gla: 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA: 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Hydroxylation;  
 KW Bone.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 9 9 HYDROXYLATION.  
 FT MOD\_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 24 24 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 23 29 BY SIMILARITY.  
 SQ SEQUENCE 49 AA; 5431 MW; 7B218B71F0312253 CRC64;  
 Query Match 73.2%; Score 199; DB 1; Length 49;  
 Best Local Similarity 76.7%; Pred. No. 1.8e-20;  
 Matches 33; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 Qy 7 GAPVYPPDPLXPRRXVXCXLPDDELADHIGFQEAAYRRFGPV 49  
 Db 7 GAPVYPPDPLXPRRXVXCXLPDDELADHIGFQEAAYRRFGPV 49  
 RESULT 9  
 OSTC\_DRONO STANDARD; PRT; 48 AA.  
 ID OSTC\_DRONO  
 AC p15504:  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (bone Gla-  
 DE protein) (BGP).  
 GN BGLAP.  
 OS Dromas novae-hollandiae (Emu).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Palaeognathae; Casuariiformes; Dromalidae;  
 OC Dromas.  
 ON NCBI\_TaxId=8790;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=88134266; PubMed=3501719;  
 RA Hsu N.L., Tseng A., Chapman G.E.;  
 RT "The amino acid sequence of Emu osteocalcin: gas phase sequencing of  
 RT Gla-containing proteins.";  
 RL Biochem. Int. 15:271-277(1987).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 DR PIR: S02208; S02208.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00594; gla: 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA: 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone.  
 FT MOD\_RES 16 16 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 20 20 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 23 23 GAMMA-CARBOXYGLUTAMIC ACID.

FT DISULFID 22 28 BY SIMILARITY.  
 SQ SEQUENCE 48 AA; 5292 MW; 50A00F3BABC7FFD CRC64;  
 Query Match 66.5%; Score 181; DB 1; Length 48;  
 Best Local Similarity 86.5%; Pred. No. 4.8e-18;  
 Matches 32; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Qy 13 PDLXPRRXVXCXLPDDELADHIGFQEAAYRRFGPV 49  
 Db 12 PDLXPRRXVXCXLPDDELADHIGFQEAAYRRFGPV 48  
 RESULT 10  
 OSTC\_CHICK STANDARD; PRT; 97 AA.  
 ID OSTC\_CHICK  
 AC P02822; Q90620;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)  
 DE (bone Gla-protein) (BGP).  
 GN BGLAP.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 ON NCBI\_TaxId=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White leghorn; TISSUE=Bone;  
 RX MEDLINE=9526465; PubMed=7747623;  
 RA Neugebauer B.M., Moore M.A., Broess M., Gerstenfeld L.C.,  
 RA Hauschka P.V.;  
 RT "Characterization of osteocalcin sequences in the chicken osteocalcin  
 RT gene: expression of osteocalcin by maturing osteoblasts and by  
 RT hypertrophic chondrocytes in vitro.";  
 RL J. Bone Miner. Res. 10:157-163(1995).  
 RN [2]  
 RP SEQUENCE OF 49-97.  
 RX MEDLINE=82007831; PubMed=6792200;  
 RA Carr S.A., Hauschka P.V., Bilemann K.;  
 RT "Gas chromatographic mass spectrometric sequence determination of  
 RT osteocalcin, a gamma-carboxyglutamic acid-containing protein from  
 RT chicken bone.";  
 RL J. Biol. Chem. 256:9944-9950(1981).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- PTM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 DR EMBL: U10578; AAY78809.1; .  
 DR PIR: A03305; GECH.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00594; gla: 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR SMART: SM00069; GLA: 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 KW Calcium-binding; Gamma-carboxyglutamic acid; Vitamin K; Bone; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT PROPEP 21 48  
 FT CHAIN 49 97 OSTEOCALCIN.





DR	InterPro: IPR002384; GLA_bone.
DR	InterPro: IPR000294; Vitk_dep_GLA.
DR	Pfam: PF00594; gla; 1.
DR	PRINTS: PR00002; GLABONE.
DR	SMART: SM00069; GLA; 1.
DR	PROSITE: PS00011; GLU-CARBOXYLATION; 1.
KW	Calcium-binding; Gamma-carboxyglutamic acid; vitamin K; Bone; Signal;
KW	MultiGene family.
FT	SIGNAL
FT	PROPEL
FT	CHAIN
FT	MOD_RES
FT	MOD_RES
FT	MOD_RES
FT	MOD_RES
FT	DISULFID
FT	SEQUENCE
QY	2 LYWGAGAPYPPDPLXPRRYXCXNLPCDELADHIGFOEAYRREYFG 47
Db	47 LRRYLGASVPSPDPLEPRLEICELDPACDELSTNGGLTAYRYRG 92
RESULT 14	
ID	OSTC_XIPGL STANDARD; PRT; 47 AA.
AC	P02823;
DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Osteocalcin (Gamma-carboxyglutamic acid-containing protein) (Bone Gla-
DE	protein) (BGP).
OS	Xiphias gladius (Swordfish).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC	Xiphidae; Xiphias.
OX	NCBI_TaxID=8245;
RN	[1]
RP	SEQUENCE.
RA	Price P.A., Otsuka A.S., Poser J.W.;
RT	"Comparison of gamma-carboxyglutamic acid-containing proteins from
RT	bovine and swordfish bone: primary structure and Ca++ binding.";
RL	(In) Wasserman R.H., Coriello R.A., Carafoli E., Kretsinger R.H.,
RL	MacLennan D.H., Siegel F.L. (eds.);
RL	Calcium-binding proteins and calcium function, pp.333-337,
RL	Elsevier, New York (1977).
CC	-1- FUNCTION: CONSTITUTES 1-28 OF THE TOTAL BONE PROTEIN. IT BINDS
CC	STRONGLY TO APATITE AND CALCIUM.
CC	-1- PM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K
CC	DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE
CC	BINDING OF CALCIUM.
CC	-1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN
CC	FAMILY.
CC	PIR: A03306; GEMF.
DR	InterPro: IPR002384; GLA_bone.
DR	InterPro: IPR000294; Vitk_dep_GLA.
DR	Pfam: PF00594; gla; 1.
DR	PRINTS: PR00002; GLABONE.
DR	PROSITE: PS00011; GLU-CARBOXYLATION; 1.
KW	Calcium-binding; Gamma-carboxyglutamic acid; vitamin K; Bone.
FT	MOD_RES
FT	MOD_RES
FT	MOD_RES
FT	DISULFID
QO	SEQUENCE 47 AA; 5080 MW; 8D93FE17B7B09223 CRC64;

Query Match 33.5%; Score 91; DB 1; Length 47;  
 Best Local Similarity 48.6%; Pred. No. 7.2e-06;  
 Matches 18; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

OY 13 PDLXPRRYKXUNPDCDELADHIGFOEAYRRPYGPV 49  
 DB 9 PLOLESREVCELNVSCDEMDATGIVAIYAYGPI 45

## RESULT 15

OSTC\_SPAU STANDARD; PRT; 45 AA.  
 AC P40148;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Osteocalcin (gamma-carboxylglutamic acid-containing protein) (Bone Gla-  
 protein) (BGP).  
 OS Sparus aurata (Gilthead sea bream).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Sparidae; Sparus.  
 OX NCBI\_TaxID=8175;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE-96139691; PubMed-8567186;  
 RA Cancela M., Williamson M.K., Price P.A.;  
 RT "Amino-acid sequence of bone gla protein from the African clawed toad  
 RT xenopus laevis and the fish Sparus aurata";  
 RL Int. J. Pept. Protein Res. 46:419-423(1995).  
 CC -1- FUNCTION: CONSTITUTES 1-2% OF THE TOTAL BONE PROTEIN. IT BINDS  
 CC STRONGLY TO APATITE AND CALCIUM.  
 CC -1- P-TM: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K  
 CC DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE  
 CC BINDING OF CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN  
 CC FAMILY.  
 DR InterPro: IPR002384; GLA\_bone.  
 DR InterPro: IPR000294; Vltk\_dep\_GLA.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00002; GLABONE.  
 DR PROSITE: PS00011; GLU-CARBOXYLATION; PARTIAL.  
 KW Calcium-binding; gamma-carboxylglutamic acid; vitamin K; Bone.  
 FT MOD\_RES 11 11 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 15 15 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 18 18 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFD 17 23 BY SIMILARITY.  
 SQ SEQUENCE 45 AA; 4951 MW; 5BAE6A1D0CB4A18 CRC64;

Query Match 26.1%; Score 71; DB 1; Length 45;  
 Best Local Similarity 43.3%; Pred. No. 0.0035;  
 Matches 13; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

OY 20 RYVXUNPDCDELADHIGFOEAYRRPYGPV 49  
 DB 14 REYCELNLACEHMDTEGIIAYTAYGPI 43

Search completed: December 4, 2002, 15:44:32  
 Job time : 12 secs

GenCore version 5.1.3  
(c) 1993 - 2002 Compu

protein search, using sw model

December 4, 2002, 15:34:39 ; Search time 35 Seconds  
(without alignment)

186.551 Million cell updates/sec

US-09-462-931-2

1 YLYQWLGAPVPYPDPLEPRR.....DELADHIGFQEAYRRFYGPV 49

OLIGO  
Carbons 60 0 Carbons 60 0

908470 seqs, 133250620 residues

•

hits satisfying chosen parameters: 908470

```
length: 0
length: 20000000000
```

Listing first 45 summaries

A\_Geneseq\_101002:\*

1	/SID2/gcgdata/genseq/genseqp-emb1/A11980.DAT *
2	/SID2/gcgdata/genseq/genseqp-emb1/A11981.DAT *
3	/SID2/gcgdata/genseq/genseqp-emb1/A11982.DAT *
4	/SID2/gcgdata/genseq/genseqp-emb1/A11983.DAT *
5	/SID2/gcgdata/genseq/genseqp-emb1/A11984.DAT *
6	/SID2/gcgdata/genseq/genseqp-emb1/A11985.DAT *
7	/SID2/gcgdata/genseq/genseqp-emb1/A11986.DAT *
8	/SID2/gcgdata/genseq/genseqp-emb1/A11987.DAT *
9	/SID2/gcgdata/genseq/genseqp-emb1/A11988.DAT *
10	/SID2/gcgdata/genseq/genseqp-emb1/A11989.DAT *
11	/SID2/gcgdata/genseq/genseqp-emb1/A11990.DAT *
12	/SID2/gcgdata/genseq/genseqp-emb1/A11991.DAT *
13	/SID2/gcgdata/genseq/genseqp-emb1/A11992.DAT *
14	/SID2/gcgdata/genseq/genseqp-emb1/A11993.DAT *
15	/SID2/gcgdata/genseq/genseqp-emb1/A11994.DAT *
16	/SID2/gcgdata/genseq/genseqp-emb1/A11995.DAT *
17	/SID2/gcgdata/genseq/genseqp-emb1/A11996.DAT *
18	/SID2/gcgdata/genseq/genseqp-emb1/A11997.DAT *
19	/SID2/gcgdata/genseq/genseqp-emb1/A11998.DAT *
20	/SID2/gcgdata/genseq/genseqp-emb1/A11999.DAT *
21	/SID2/gcgdata/genseq/genseqp-emb1/A12000.DAT *
22	/SID2/gcgdata/genseq/genseqp-emb1/A12001.DAT *
23	/SID2/gcgdata/genseq/genseqp-emb1/A12002.DAT *

is the number of results predicted by chance to have a higher than or equal to the score of the result being printed, derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49	100.0	49	20	AAW9811	Human osteocalcin
2	49	100.0	51	12	AAR10146	Human osteocalcin
3	49	100.0	98	19	AAW76094	Human osteocalcin
4	49	100.0	100	12	AAR10147	Human osteocalcin
5	49	100.0	100	23	AAU10687	Reference sequence
6	43	87.8	97	13	AAK20043	Fusion protein for
7	30	61.2	100	23	AAU10688	Polymorphic variant
8	28	57.1	49	18	AAW34263	Glu1-osteocalcin
9	28	57.1	49	18	AAW34265	Glu21-osteocalcin
10	25	51.0	46	22	AAW91919	Bone Gla protein F

1.1	25	51.0	47	22	AA819117
1.2	25	51.0	49	14	AA832936
1.3	25	51.0	49	14	AA832937
1.4	25	51.0	49	18	AA834266
1.5	25	51.0	49	18	AA834267
1.6	20	40.8	21	11	AA806630
1.7	20	40.8	16	16	AA806631
1.8	19	38.8	20	15	AA801681
1.9	14	28.6	15	11	AA806633
2.0	14	28.6	68	23	AA801784
2.1	13	26.5	13	18	AA804605
2.2	13	26.5	13	21	AA858430
2.3	13	26.5	13	22	AA856568
2.4	13	26.5	13	22	AA802088
2.5	13	26.5	13	22	AA819202
2.6	13	26.5	13	22	AA801727
2.7	13	26.5	13	23	AA815409
2.8	13	26.5	13	23	AA883455
2.9	13	26.5	13	23	AA883572
3.0	13	26.5	13	23	AA878634
3.1	13	26.5	13	23	AA818828
3.2	13	26.5	13	23	AA818839
3.3	13	26.5	13	23	AA848056
3.4	13	26.5	15	21	AA870693
3.5	13	26.5	18	20	AA805366
3.6	13	26.5	38	23	AA888451
3.7	13	26.5	39	23	AA888442
3.8	13	26.5	43	18	AA818997
3.9	13	26.5	44	23	AA884388
4.0	13	26.5	44	23	AA888447
4.1	13	26.5	45	23	AA888443
4.2	13	26.5	46	23	AA888440
4.3	13	26.5	55	23	AA888497
4.4	13	26.5	55	23	AA888497
4.5	13	26.5	58	23	AA888455

## ALIGNMENTS

XX	AAW99811	AAW99811 standard; Protein; 49 AA.
XX	AAW99811;	
XX	08-JUN-1999	(first entry)
XX	Human osteocalcin	fragment.
XX	Human; osteocalcin; urine; gamma-carboxylated; bone	
KW	bone turnover; puberty; menopause; detection.	
XX	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	Modified-site	17
FT	Modified-site	/note= "possibly gamma-carboxylated"
FT	Modified-site	21
FT	Modified-site	/note= "possibly gamma-carboxylated"
FT	Modified-site	24
XX	W09909058-A1.	/note= "possibly gamma-carboxylated"
XX	25-FEB-1999.	
XX	24-JUN-1998;	98W0-FT00550.
PR	15-AUG-1997;	97FI-0003371.
PA	(HELLA) HELLMAN J.	
PA	(KAERK/) KAERKONEN S.	

Bone Gla protein P  
21, 24, gla human  
17, 21, 24, gla hu  
Gla21-osteocalcin  
Human osteocalcin  
Bone Gla protein.  
Osteocalcin AA 1-1.  
Human osteocalcin  
Peptide presentation  
Osteocalcin 7-19 f  
E tag peptide, SEC  
E tag used for rec  
E-tag peptide. Syn  
Bone Gla protein p  
E tag. Synthetic  
E tag - peptide ra  
E tag. Synthetic.  
Synthetic E tag pe  
Synthetic E tag pe  
E tag used in the  
Human BGP peptide  
E-tag sequence. S  
E. coli antibody d  
HBV specific singl  
Insulin/insulin-11  
Insulin/insulin-11  
Vector PC019EHIS  
Insulin/insulin-11  
Insulin/insulin-11  
Insulin/insulin-11  
Insulin/insulin-11  
Insulin/insulin-11  
Insulin/insulin-11

PA (KARP/) KARP M.  
 PA (LOEV/) LOEVGREN T.  
 PA (PETT/) PETERSSON K.  
 PA (VAEA/) VAEAEVAENEN H K.  
 XX  
 PI Hellman J, Kaekonen S, Karp M, Loevgren T, Pettersson K;  
 PI Vaeseenaenen HK;  
 DR WPI: 1999-180972/15.  
 DR N-PSDB: AAX19767.  
 XX  
 PT Human osteocalcin peptide fragments - useful for detecting  
 PT metabolic bone turnover rates and disorders  
 XX  
 PS Claim 1; Fig 1A; 49pp; English.  
 CC An assay has been developed for the detection of gamma-carboxylated  
 CC human osteocalcin found in urine, where the glutamic acid residues at  
 CC positions 17, 21 and 24 are gamma-carboxylated. The assay is used to  
 CC measure the rate of bone turnover (formation and/or resorption) and/or  
 CC for investigating metabolic bone disorders in individuals. Especially  
 CC mentioned, the assays can detect differences between adults and children  
 CC going through puberty, pre- and post-menopausal women and children  
 CC having high bone turnover. The assay allows more sensitive detection of  
 CC human osteocalcin (hOC) in urine than in serum. Significant increases in  
 CC hOC (40-48%) were detected in serum in menopause (normal increase is  
 CC 30-50% above pre-menopausal women). Detection of urine hOC using the  
 CC assay showed an increase as high as 75-79%. The present sequence  
 CC represents a fragment of hOC.  
 CC  
 SQ Sequence 49 AA:

Query Match 100.0%; Score 49; DB 20; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-45;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLPRREVCELNPCDELADHIGFOEAYRRFGPV 49  
 ||||||||||||||||||||||||||||||||||||||||||||  
 DB 1 YLYQWLGAPVPPDPLPRREVCELNPCDELADHIGFOEAYRRFGPV 49

RESULT 2  
 AAR10146  
 ID AAR10146 standard; peptide; 51 AA.  
 XX  
 AC AAR10146;  
 XX  
 DT 27-MAR-1991 (first entry)  
 XX  
 DE Human osteocalcin precursor polypeptide (I).  
 XX  
 KW Human osteocalcin precursor polypeptide; OC; carboxypeptidase B;  
 KW calcium; vitamin K; bone formation; dysbolism.  
 XX  
 OS Homo sapiens.

Key Location/Qualifiers  
 FH Misc-difference 1..1  
 FT Misc-difference /note="residue may be omitted"  
 FT Misc-difference 51..51  
 FT Misc-difference /note="residue may be omitted"

JP02201294-A.  
 XX  
 PD 03-DEC-1990.  
 XX  
 PF 22-AUG-1989; 89JP-0214239.  
 XX  
 PR 06-DEC-1988; 88JP-0306931.  
 PR 22-AUG-1989; 89JP-0214239.  
 XX  
 PA (TAKA-) TAKARA SYUZO KK.  
 XX

DR WPI: 1991-018865/03.  
 XX  
 PT Human osteocalcin (OC) precursor polypeptide - used to prepare  
 PT purified OC by digestion with carboxypeptidase B.  
 XX  
 PS Claim 1; Page 1; 11pp; Japanese.  
 XX  
 CC To K1 is attached H and to K51 is attached OH.  
 CC A novel gene encoding human OC precursor polymer was inserted into  
 CC plasmid pOC 980. E. coli HB101 was transformed with this plasmid, and  
 CC made to efficiently produce the objective polymer. The polymer was  
 CC treated with lysylendopeptidase B to remove spacers. The obtained  
 CC human precursor was further treated with carboxypeptidase to prepare a  
 CC purified human OC.  
 CC OC is a calcium bonded protein (depending on vitamin K) produced in the  
 CC bone. It is thought to be a promoting factor during the bone  
 CC formation and may be used to treat diseases due to dysbolism of the  
 CC bone.  
 CC See also AAR10147 and AAQ10193-98.  
 CC  
 SQ Sequence 51 AA:

Query Match 100.0%; Score 49; DB 12; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-45;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVPPDPLPRREVCELNPCDELADHIGFOEAYRRFGPV 49  
 ||||||||||||||||||||||||||||||||||||||||||||  
 DB 2 YLYQWLGAPVPPDPLPRREVCELNPCDELADHIGFOEAYRRFGPV 50

RESULT 3  
 AAW76094  
 ID AAW76094 standard; Protein; 98 AA.  
 XX  
 AC AAW76094;  
 XX  
 DT 21-DEC-1998 (first entry)  
 XX  
 DE Human osteocalcin protein.

XX Promoter; tissue-specific gene expression; skeletal tissue;  
 KW stem cell; bone; cartilage; osteocalcin; hOC gene; human;  
 KW osteoporosis; osteopenia; osteosarcoma; cancer; metastasis;  
 KW gene therapy.  
 XX

OS Homo sapiens.

XX W09839427-A2.

PD 11-SEP-1998.

PF 06-MAR-1998; 98MO-US04421.

PR 06-MAR-1997; 97US-0039839.

PA (UYMA-) UNTV MASSACHUSETTS.

PI Frenkel B, Hou Z, Lian JB, Nilsson S, Peters S;  
 PI Quesenberry P, Stein GS, Stein JL;

DR WPI: 1998-495839/42.

DR N-PSDB: AAV46429.

PT Expression of exogenous genes in differentiated cells - by  
 PT transducing pluripotent stem cells capable of maturing into  
 PT differentiated cells with nucleic acid comprising exogenous gene,  
 PT useful for, e.g. treatment of osteoporosis

PS Disclosure; Page 33-34; 63pp; English.

CC This is the amino acid sequence of the human osteocalcin, deduced  
 CC from the coding exons of the human hOC gene (see AAV46429). The

CC invention pertains to a method for expressing endogenous genes in  
CC differentiated cells of a specific type. The method involves  
CC contacting pluripotent stem cells capable of maturing into  
CC differentiated cells with a nucleic acid comprising an exogenous  
CC gene linked to a regulatory element capable of controlling expression  
CC of the exogenous gene in the differentiated cells. A population of  
CC transduced stem cells capable of maturing into differentiated cells  
CC expressing the exogenous gene is produced. Preferably, the  
CC differentiated cells are in a tissue of interest, such as bone or  
CC cartilage, and the exogenous gene is operably linked to at least one  
CC osteocalcin regulatory element derived from the hOC promoter. The  
CC exogenous gene can encode a therapeutic protein useful for treating  
CC a disease, especially osteoporosis, osteopenia, osteosarcoma,  
CC primary malignancy or metastases (all claimed).  
CC  
SQ Sequence 98 AA;  
  
Query Match 100.0%; Score 49; DB 19; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.9e-45;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 YLYWMLGAPVPPPLEPRREVCELPDCLADHIGFOEAYRRFGPV 49  
|||||  
50 YLYWMLGAPVPPPLEPRREVCELPDCLADHIGFOEAYRRFGPV 98  
Db  
  
RESULT 4  
AAR10147  
ID AAR10147 standard; Protein: 100 AA.  
XX  
XX AAR10147;  
AC  
XX  
XX 27-MAR-1991 (first entry)  
DT  
XX  
DE Human osteocarcin precursor polypeptide (II).  
XX  
XX Human osteocarcin precursor polypeptide; OC: carboxypeptidase B;  
KW calcium; vitamin K; bone formation; dysbolism.  
XX  
OS Homo sapiens.  
XX  
XX JF02201294-A.  
PN  
XX  
PD 03-DEC-1990.  
XX  
XX 22-AUG-1989; 89JP-0214239.  
PF  
XX  
XX 06-DEC-1988; 88JP-0306931.  
PR  
XX 22-AUG-1989; 89JP-0214239.  
PA  
XX (TAKA-) TAKARA SYUZO KK.  
XX  
XX WPI: 1991-018865/03.  
DR  
XX  
XX Human osteocarcin (OC) precursor polypeptide - used to prepare  
PT purified OC by digestion with carboxypeptidase B.  
PT  
XX  
XX Claim 5; Page 1; 11pp; Japanese.  
PS  
XX  
XX To K1 is attached H and to V100 is attached OH.  
CC A novel gene encoding human OC precursor polymer was inserted into  
CC plasmid pOC 980. E.coli HB101 was transformed with this plasmid, and  
CC made to efficiently produce the objective polymer. The polymer was  
CC treated with lysine endopeptidase B to remove spacers. The obtained  
CC human precursor was further treated with carboxy peptidase to prepare a  
CC purified human OC.  
CC OC is a calcium bonded protein (depending on vitamin K) produced in the  
CC bone. It is thought to be a promoting factor during the bone  
CC formation and may be used to treat diseases due to dysbolism of the  
CC bone.  
CC See also AAR10147 and AAQ10193-98.  
CC  
XX  
XX Sequence 100 AA;

Query Match 100.0%; Score 49; DB 12; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3e-45;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 YLYWMLGAPVPPPLEPRREVCELPDCLADHIGFOEAYRRFGPV 49  
|||||  
52 YLYWMLGAPVPPPLEPRREVCELPDCLADHIGFOEAYRRFGPV 100  
Db  
  
RESULT 5  
AAU10687  
ID AAU10687 standard; Protein: 100 AA.  
XX  
XX AAU10687;  
AC  
XX  
XX 14-FEB-2002 (first entry)  
DT  
XX  
XX Reference sequence for human BGLAP protein.  
DE  
XX  
XX Human: single nucleotide polymorphism; SNP: BGLAP; chromosome 1q25-q31;  
KW bone gamma carboxyglutamate protein; haplotyping; genotyping;  
KW osteoporosis; osteopathic.  
XX  
XX  
XX Homo sapiens.  
XX  
XX W0200177131-A2.  
PN  
XX  
XX 18-OCT-2001.  
PD  
XX  
XX 11-APR-2001; 2001WO-US12010.  
PF  
XX  
XX 11-APR-2000; 2000US-195840P.  
PR  
XX  
XX (GENA-) GENAISSANCE PHARM INC.  
PA  
XX  
XX Bentivegna SC, Chew A, Chol JY, Koshy B, Rounds E, Stephens JC;  
PI  
XX  
XX WPI: 2002-041288/05.  
DR  
XX N-PSDB; AAS16363, AAS16364.  
XX  
XX  
XX New haplotypes of the human bone gamma carboxyglutamate protein gene,  
PT useful to diagnose and treat diseases associated with the gene such as  
PT osteoporosis  
PT  
XX  
XX Claim 27; Fig 3; 53pp; English.  
PS  
XX  
XX The present invention relates to novel single nucleotide polymorphisms  
CC (SNPs) in the human bone gamma carboxyglutamate protein (BGLAP) gene  
CC located on chromosome 1q25-q31, and methods for haplotyping and/or  
CC genotyping the BGLAP gene in an individual. The methods of the  
CC invention make use of allele-specific oligonucleotides (ASOs) as probes  
CC and primers and/or primer-extension oligonucleotides for detecting the  
CC BGLAP gene polymorphisms. The polymucleotides and screened compounds are  
CC useful for (developing) treatment of diseases associated with BGLAP  
CC activity, such as osteoporosis. The present sequence represents a  
CC reference sequence for the BGLAP protein.  
CC  
XX  
XX Sequence 100 AA;  
  
Query Match 100.0%; Score 49; DB 23; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3e-45;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 YLYWMLGAPVPPPLEPRREVCELPDCLADHIGFOEAYRRFGPV 49  
|||||  
52 YLYWMLGAPVPPPLEPRREVCELPDCLADHIGFOEAYRRFGPV 100  
Db  
  
RESULT 6  
AAR20043  
ID AAR20043 standard; Protein: 97 AA.  
XX

```

AC AAR20043;
XX
XX 09-APR-1992 (first entry)
XX
DE Fusion protein for expression of human osteocalcin.
XX
XX gamma-carboxyglutamic acid; bone matrix; Gla protein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..22 /label= signal_peptide
FT Peptide 23..44 /label= pro_peptide
FT Cleavage-site 45..48
FT Protein 49..97 /label= osteocalcin
FT
XX EP463571-A.
XX
XX 02-JAN-1992.
XX
XX 20-JUN-1991; 91EP-0110173.
XX
XX 30-NOV-1990; 90JP-0330146.
XX PR 20-JUN-1990; 90JP-0159909.
XX
XX (TEIJ ) TEIJIN KK.
XX
XX Eguchi H, Kamimura TF, Sugiyama T, Hosoda K;
XX PI WPI; 1992-009183/02.
XX DR N-PSDB; AAG20210.
XX
XX Human osteocalcin prodn. - using DNA coding for human osteocalcin
PT fusion protein for expression in host cells
XX
XX Claim 15; Fig 7; 53pp; English.
XX
CC This sequence is a specific example of a claimed generic fusion
CC protein comprising human osteocalcin. The pro-peptide is recognised
CC by an enzyme capable of Glu to Gla conversion on human osteocalcin.
CC The recombinant protein was obtained by culturing host cells
CC transformed with a vector containing the synthetic coding sequence.
CC The Glu residues could then be converted to Gla (i.e.
CC gamma-carboxyglutamic acid) and the osteocalcin sequence cleaved
CC from the propeptide. The mature protein is suitable for use in
CC immunoassays and as a drug for treatment of bone metabolism
CC disorders. See also AAR20044-6.
XX
XX
SQ Sequence 97 AA;
XX
Query Match 87.8%; Score 43; DB 13; Length 97;
Best Local Similarity 100.0%; Pred. No. 8.7e-39;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 YLYQWLGAIPYPPDPLEPRREVCENLPDDELADHIGFOEAYR 43
DB 49 YLYQWLGAIPYPPDPLEPRREVCENLPDDELADHIGFOEAYR 91
XX
RESULT 7
AAU10688
ID AAU10688 standard; Protein; 100 AA.
XX
XX AAU10688;
XX
XX 14-FEB-2002 (first entry)
XX
XX Polymorphic variant of human BGLAP protein.
DE
XX Human; single nucleotide polymorphism; SNP; BGLAP; chromosome 1q25-q31;

```

```

KW bone gamma carboxyglutamate protein; haplotyping; genotyping;
XX osteoporosis; osteopathic; variant.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 82 /note= "Substitution of Glu to Lys""
FT FT Misc-difference 94 /note= "Substitution of Arg to Gln"
FT
XX WO200177131-A2.
XX
XX 18-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US12010.
XX PE 11-APR-2000; 2000US-195840P.
XX PR
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Bentivegna SC, Chew A, Choi JY, Koshy B, Rounds E, Stephens JC;
XX PI WPI; 2002-041288/05.
XX
XX New haplotypes of the human bone gamma carboxyglutamate protein gene,
PT useful to diagnose and treat diseases associated with the gene such as
PT osteoporosis
XX
XX Claim 27; Page -: 53pp; English.
XX
XX
XX The present invention relates to novel single nucleotide polymorphisms
CC (SNPs) in the human bone gamma carboxyglutamate protein (BGLAP) gene
CC located on chromosome 1q25-q31, and methods for haplotyping and/or
CC genotyping the BGLAP gene in an individual. The methods of the
CC invention make use of allele-specific oligonucleotides (ASOs) as probes
CC and primers and/or primer-extension oligonucleotides for detecting the
CC BGLAP gene polymorphisms. The polynucleotides and screened compounds are
CC useful for (developing) treatment of diseases associated with BGLAP
CC activity, such as osteoporosis. The present sequence represents a
CC polymorphic variant of the BGLAP protein (AAU10687).
CC Note: The present sequence is not given in the specification but is
CC created by the indexer from the information given in the patent.
XX
XX
SQ Sequence 100 AA;
XX
Query Match 61.2%; Score 30; DB 23; Length 100;
Best Local Similarity 100.0%; Pred. No. 9.5e-25;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 YLYQWLGAIPYPPDPLEPRREVCENLPDCC 30
DB 52 YLYQWLGAIPYPPDPLEPRREVCENLPDCC 81
XX
RESULT 8
AAW34263
ID AAW34263 standard; peptide; 49 AA.
XX
XX AAW34263;
XX
XX 23-APR-1998 (first entry)
XX
XX Glu17-osteocalcin peptide #1.
XX
XX Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;
KW osteoporosis; human.
XX
XX Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 21 /note= "gamma-carboxyglutamic acid"

```

XX WO9738309-A1.  
XX 16-OCT-1997.  
XX 10-APR-1997; 97WO-JP01246.  
XX 27-FEB-1997; 97JP-0043331.  
XX 10-APR-1996; 96JP-0088608.  
XX (EISA ) EISAI CO LTD.  
XX Kimura T, Morimoto S, Sakakibara S;  
XX WPI; 1997-512875/47.  
XX Antibody specific for Glu17-osteocalcin, or its fragment - for  
XX diagnosis of bone disorders such as osteoporosis  
XX  
XX Claim 3; Page 16; 28pp: Japanese.  
XX  
XX This sequence represents a human Glu17-osteocalcin peptide. This sequence  
XX is recognised by the antibody of the invention. The antibody of the  
XX invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
XX binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The  
XX antibody can be used for the diagnosis of bone related disorders, such as  
XX osteoporosis.  
XX  
XX Sequence 49 AA:  
SQ  
Query Match 57.1%; Score 28; DB 18; Length 49;  
Best Local Similarity 100.0%; Pred. No. 7.4e-23;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 22 VCELNPPDCDELADHIGFQEAARYRFGPY 49  
Db 22 VCELNPPDCDELADHIGFQEAARYRFGPY 49  
RESULT 9  
AAW34265  
ID AAW34265 standard; peptide; 49 AA.  
XX  
XX AAW34265;  
AC  
XX 23-APR-1998 (first entry)  
DT  
XX  
XX Gla21-osteocalcin peptide #1.  
DE  
XX  
XX Gla17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
KW osteoporosis; human.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH Modified-site 17  
FT /note- "gamma-carboxyglutamic acid"  
FT Modified-site 21  
FT /note- "gamma-carboxyglutamic acid"  
PN WO9738309-A1.  
XX  
XX 16-OCT-1997.  
PD  
XX  
XX 10-APR-1997; 97WO-JP01246.  
XX  
XX 27-FEB-1997; 97JP-0043331.  
PR  
XX 10-APR-1996; 96JP-0088608.  
XX  
XX (EISA ) EISAI CO LTD.  
XX  
XX Kimura T, Morimoto S, Sakakibara S;  
PI  
XX

DR WPI; 1997-512875/47.  
XX  
XX Antibody specific for Glu17-osteocalcin, or its fragment - for  
XX diagnosis of bone disorders such as osteoporosis  
XX  
XX Claim 4; Page 17; 28pp: Japanese.  
XX  
XX This sequence represents a human Gla21-osteocalcin peptide. This sequence  
XX is recognised by the antibody of the invention. The antibody of the  
XX invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
XX binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The  
XX antibody can be used for the diagnosis of bone related disorders, such as  
XX osteoporosis.  
XX  
XX Sequence 49 AA:  
SQ  
Query Match 57.1%; Score 28; DB 18; Length 49;  
Best Local Similarity 100.0%; Pred. No. 7.4e-23;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 22 VCELNPPDCDELADHIGFQEAARYRFGPY 49  
Db 22 VCELNPPDCDELADHIGFQEAARYRFGPY 49  
RESULT 10  
AAB91919  
ID AAB91919 standard; Peptide; 46 AA.  
XX  
XX AAB91919;  
AC  
XX 22-JUN-2001 (first entry)  
DT  
XX  
XX Bone Gla protein peptide SEQ ID NO:1095.  
DE  
XX  
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimideyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Synthetic.  
OS  
XX  
XX WO200069900-A2.  
PN  
XX  
XX 23-NOV-2000.  
PD  
XX  
XX 17-MAY-2000; 2000WO-US13576.  
PF  
XX  
XX 17-MAY-1999; 99US-0134406.  
PR  
XX 10-SEP-1999; 99US-0153406.  
PR  
XX 15-OCT-1999; 99US-0159783.  
XX  
XX (CONF-) CONJUCHEM INC.  
XX  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
PI  
XX WPI; 2001-112059/12.  
DR  
XX  
XX Modifying and attaching therapeutic peptides to albumin prevents  
XX peptidase degradation, useful for increasing length of in vivo activity  
XX  
XX Disclosure; Page 553; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
XX comprising a therapeutically active amino acid region (III) and a  
XX reactive group (II) (e.g. succinimideyl and maleimido groups) attached to  
XX a less therapeutically active amino acid region (IV), which covalently  
XX bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX factors and neurotransmitters, to protect them from peptidase activity  
XX in vivo for the treatment of various disorders. Endogenous therapeutic

CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.

SO Sequence 46 AA;

Query Match 51.0%; Score 25; DB 22; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
 ||||||||||||||||||||  
 DB 22 LNPDCDELADHIGFOEAYRRFGPV 46

RESULT 11  
 AAB91917  
 ID AAB91917 standard; peptide; 47 AA.  
 AC AAB91917;  
 DT 22-JUN-2001 (first entry)  
 XX  
 DE Bone Gla protein peptide SEQ ID NO:1093.  
 XX  
 KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidyl; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO200069900-A2.  
 PD 23-NOV-2000.  
 XX  
 PF 17-MAY-2000; 2000MO-US13576.  
 XX  
 PR 17-MAY-1999; 99US-0134406.  
 PR 10-SEP-1999; 99US-0153406.  
 PR 15-OCT-1999; 99US-0159783.  
 XX  
 PA (CONF-) CONJUCHEM INC.  
 XX  
 PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
 XX  
 DR WPI; 2001-112059/12.  
 XX  
 PT Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity  
 XX  
 PS Disclosure; Page 552-553; 733pp; English.  
 XX  
 CC The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases

CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.

SO Sequence 47 AA;

Query Match 51.0%; Score 25; DB 22; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
 ||||||||||||||||||||  
 DB 23 LNPDCDELADHIGFOEAYRRFGPV 47

RESULT 12  
 AAR32936  
 ID AAR32936 standard; peptide; 49 AA.  
 AC AAR32936;  
 DT 05-JUL-1993 (first entry)  
 XX  
 DE 21, 24, Gla human osteocalcin peptide.  
 XX  
 KW Gamma-carboxyglutamic acid; OS.  
 XX  
 OS Synthetic.  
 XX  
 FH key Location/Qualifiers  
 FT Misc-difference 21  
 FT /label= OTHER  
 FT /note= "OTHER= gamma-carboxyglutamic acid"  
 FT Misc-difference 24  
 FT /label= OTHER  
 FT /note= "OTHER= gamma-carboxyglutamic acid"  
 FT  
 FT  
 PN JP05032697-A.  
 XX  
 PD 09-FEB-1993.  
 XX  
 PF 31-JUL-1991; 91JP-0213251.  
 XX  
 PR 31-JUL-1991; 91JP-0213251.  
 XX  
 PA (TEIJ ) TEIJIN LTD.  
 XX  
 DR WPI; 1993-088665/11.  
 XX  
 PT Synthetic human osteocalcin for standard in determ. of natural  
 PT osteocalcin - prepd. by introducing gamma-carboxyglutamic acid  
 PT as fluorenyl protected gp.  
 XX  
 PS Claim 1; Page 2; 10pp; Japanese.  
 XX  
 CC The synthetic 21, 24, gamma-carboxyglutamic acid form of human  
 CC osteocalcin (OS) was produced by introducing protected Gla. The  
 CC substance may be produced in high yield and is useful as standard  
 CC for the determination of human OS.  
 CC See also AAR32937.  
 XX  
 SO Sequence 49 AA;

Query Match 51.0%; Score 25; DB 14; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-19;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFGPV 49  
 ||||||||||||||||||||  
 DB 25 LNPDCDELADHIGFOEAYRRFGPV 49

RESULT 13



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:32:49 : Search time 14 Seconds  
(without alignments)  
102.980 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 284  
Sequence: 1 YLYQWLGAPVPPDPLEPRR.....DELADHIGQEAATRRFGPV 49

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	100.0	49	6	5434245-1 Patent No. 5434245
2	284	100.0	50	6	5434245-2 Patent No. 5434245
3	284	100.0	50	6	5434245-3 Patent No. 5434245
4	284	100.0	51	6	5434245-4 Patent No. 5434245
5	272	95.8	49	6	5164483-2 Patent No. 516448
6	266	93.7	49	6	5164483-1 Patent No. 516448
7	266	93.7	98	6	5168041-3 Patent No. 5168041
8	248.5	87.5	48	6	5168041-1 Patent No. 5168041
9	76	26.8	13	2	US-08-796-598-13 Sequence 13, App1
10	76	26.8	13	2	US-08-447-175A-13 Sequence 13, App1
11	76	26.8	13	2	US-08-943-915-6 Sequence 6, App1
12	76	26.8	13	3	US-08-881-037-112 Sequence 112, App1
13	76	26.8	13	3	US-08-652-816A-43 Sequence 43, App1
14	68	23.9	13	4	US-09-142-974B-5 Sequence 5, App1
15	63	22.2	2254	2	US-08-677-010-3 Sequence 3, App1
16	63	22.2	2254	2	US-08-790-519-3 Sequence 3, App1
17	60	21.1	353	2	US-09-068-109-2 Sequence 2, App1
18	58.5	20.6	12	3	US-08-968-747-7 Sequence 7, App1
19	58.5	20.6	12	3	US-08-493-071-28 Sequence 28, App1
20	57.5	20.2	436	4	US-09-734-673-2 Sequence 2, App1
21	57.5	20.2	436	4	US-09-523-849-2 Sequence 2, App1
22	57	20.1	765	4	US-08-737-109-11 Sequence 11, App1
23	55	19.4	419	4	US-09-071-035-324 Sequence 324, App1
24	55	19.4	450	4	US-09-071-035-322 Sequence 322, App1
25	54.5	19.2	486	1	US-07-672-483-2 Sequence 2, App1
26	54	19.0	106	3	US-09-083-351-7 Sequence 7, App1
27	54	19.0	106	4	US-09-083-352-7 Sequence 7, App1

## ALIGNMENTS

28	53.5	18.8	423	1	US-08-523-376-3	Sequence 3, App1
29	53.5	18.8	466	4	US-08-259-451-13	Sequence 13, App1
30	53	18.7	666	4	US-08-937-067-10	Sequence 10, App1
31	53	18.7	1956	4	US-08-843-417-2	Sequence 2, App1
32	53	18.7	1957	4	US-08-669-656A-8	Sequence 8, App1
33	53	18.7	2132	4	US-08-669-656A-6	Sequence 6, App1
34	52	18.3	11	6	5168041-2	Patent No. 5168041
35	52	18.3	1957	4	US-08-669-656A-2	Sequence 2, App1
36	51	18.0	323	2	US-08-435-149-2	Sequence 2, App1
37	51	18.0	324	1	US-08-310-416A-14	Sequence 14, App1
38	51	18.0	324	2	US-08-888-171-14	Sequence 14, App1
39	51	18.0	431	4	US-09-038-832-2	Sequence 2, App1
40	51	18.0	431	4	US-09-038-832-4	Sequence 4, App1
41	51	18.0	577	2	US-08-435-149-3	Sequence 3, App1
42	51	18.0	585	4	US-08-937-067-9	Sequence 9, App1
43	51	18.0	611	4	US-09-475-460A-32	Sequence 32, App1
44	51	18.0	611	4	US-09-748-061A-32	Sequence 32, App1
45	51	18.0	614	1	US-08-652-207A-2	Sequence 2, App1

RESULT 1  
5434245-1  
; Patent No. 5434245  
; APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION NUMBER: US/07/993,980  
; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989  
; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989  
; SEQ ID NO: 1:  
; LENGTH: 49  
5434245-1  
Query Match 100.0%; Score 284; DB 6; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.2e-30;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 YLYQWLGAPVPPDPLEPRREVCLELNPDCDELADHIGFOEAATRRFGPV 49  
DB 1 YLYQWLGAPVPPDPLEPRREVCLELNPDCDELADHIGFOEAATRRFGPV 49  
RESULT 2  
5434245-2  
; Patent No. 5434245  
; APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; FILING DATE: 16-DEC-1992  
; PRIOR APPLICATION NUMBER: 855,473; 444,786  
; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989  
; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989  
; SEQ ID NO: 2:  
; LENGTH: 50  
5434245-2  
Query Match 100.0%; Score 284; DB 6; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.3e-30;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49  
1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49

RESULT 3  
5434245-3

; Patent No. 5434245  
; APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME

; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/993,980  
; FILING DATE: 16-DEC-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989

; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989

; SEQ ID NO:3;  
; LENGTH: 50

5434245-3

Query Match 100.0%; Score 284; DB 6; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.3e-30;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49  
2 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 50

RESULT 4

5434245-4

; Patent No. 5434245  
; APPLICANT: KOYAMA, NOBUTO; KIMIZUKA, FUSAO; KATO, IKUNOSHIN  
; TITLE OF INVENTION: POLYPEPTIDES AND METHOD FOR PREPARING  
; THE SAME

; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/993,980  
; FILING DATE: 16-DEC-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 855,473; 444,786  
; FILING DATE: 23-MAR-1992; 01-DEC-1989

; APPLICATION NUMBER: 444,786  
; FILING DATE: 01-DEC-1989

; SEQ ID NO:4;  
; LENGTH: 51

5434245-4

Query Match 100.0%; Score 284; DB 6; Length 51;  
Best Local Similarity 100.0%; Pred. No. 1.3e-30;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49  
2 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 50

RESULT 5

5164483-2

; Patent No. 5164483  
; APPLICANT: TAKASHI, KURIHARA, EIJI, TANIYAMA, SACHIO, HIROSE  
; TITLE OF INVENTION: Y-CARBOXYGLUTAMATE DERIVATIVE METHOD  
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN  
; USING THE SAME

; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/575,639  
; FILING DATE: 31-AUG-1990

; SEQ ID NO:2;  
; LENGTH: 49

Query Match 95.8%; Score 272; DB 6; Length 49;  
Best Local Similarity 95.9%; Pred. No. 4.7e-29;  
Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49  
1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49

RESULT 6

5164483-1

; Patent No. 5164483  
; APPLICANT: TAKASHI, KURIHARA, EIJI, TANIYAMA, SACHIO, HIROSE  
; TITLE OF INVENTION: Y-CARBOXYGLUTAMATE DERIVATIVE METHOD  
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN  
; USING THE SAME

; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/575,639  
; FILING DATE: 31-AUG-1990

; SEQ ID NO:1;  
; LENGTH: 49

5164483-1

Query Match 93.7%; Score 266; DB 6; Length 49;  
Best Local Similarity 93.9%; Pred. No. 2.9e-28;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49  
1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49

RESULT 7

5164483-3

; Patent No. 5164483  
; APPLICANT: TAKASHI, KURIHARA, EIJI, TANIYAMA, SACHIO, HIROSE  
; TITLE OF INVENTION: Y-CARBOXYGLUTAMATE DERIVATIVE METHOD  
; FOR PREPARING THE SAME AND METHOD FOR PREPARING HUMAN OSTEOCALCIN  
; USING THE SAME

; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/575,639  
; FILING DATE: 31-AUG-1990

; SEQ ID NO:3;  
; LENGTH: 110

5164483-3

Query Match 93.7%; Score 266; DB 6; Length 98;  
Best Local Similarity 93.9%; Pred. No. 6.8e-28;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49  
1 YLYQWLGAVPPYPPDLEPRREVCEINPCDDELADHIGFOEAYRRFGPV 49

RESULT 8

5168041-1

; Patent No. 5168041  
; APPLICANT: BERGMANN, ANDREAS E.  
; TITLE OF INVENTION: METHOD FOR THE DETERMINATION OF  
; OSTEOCALCIN IN HUMAN SERUM OR PLASMA

; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/416,728  
; FILING DATE: 03-OCT-1989

; SEQ ID NO:1;  
; LENGTH: 48

5168041-1

Query Match 87.5%; Score 248.5; DB 6; Length 48;  
Best Local Similarity 91.8%; Pred. No. 5.8e-26;  
Matches 45; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 1 YLXQWLGAPVPPDPLEPRRVCCLNPPDCDELADHIGFOEAYRRFGPV 49  
1 YLXQWLGAPVPPDPLEPR-RXVCXLPNPPDCDELADHIGFOEAYRRFGPV 48

RESULT 9  
US-08-796-598-13  
Sequence 13, Application US/08796598  
Patent No. 5827659

GENERAL INFORMATION:  
APPLICANT: PATTERSON, DALE H.  
TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING  
TITLE OF INVENTION: POLYMERS USING MASS SPECTROMETRY.  
NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:  
ADDRESS: Patent Administrator - Testa, Hurwitz &  
ADDRESS: Thibault  
STREET: High Street Tower, 125 High Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/796,598  
FILING DATE: 07-FEB-1997  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/446,055  
FILING DATE: 19-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: FLYNN Esq., Kerry A.  
REGISTRATION NUMBER: 33,693  
REFERENCE/DOCKET NUMBER: STP-115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-796-598-13

Query Match 26.8%; Score 76; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00063;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPPDPLEPR 19  
1 GAPVPPDPLEPR 13

RESULT 10  
US-08-447-175A-13  
Sequence 13, Application US/08447175A  
Patent No. 5869240

GENERAL INFORMATION:  
APPLICANT: PATTERSON, DALE H.  
TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING

TITLE OF INVENTION: POLYMERS WITH A STATISTICAL CERTAINTY USING MASS  
TITLE OF INVENTION: SPECTROMETRY.

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESS: Patent Administrator - Testa, Hurwitz &  
ADDRESS: Thibault, LLP  
STREET: High Street Tower, 125 High Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/447,175A  
FILING DATE: 19-MAY-1995

CLASSIFICATION: 422

ATTORNEY/AGENT INFORMATION:

NAME: RAUSCHENBACH, Kurt  
REGISTRATION NUMBER: 40,137  
REFERENCE/DOCKET NUMBER: STP-114

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-447-175A-13

Query Match 26.8%; Score 76; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00063;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPPDPLEPR 19  
1 GAPVPPDPLEPR 13

RESULT 11

US-08-943-915-6  
Sequence 6, Application US/08943915  
Patent No. 5998170

GENERAL INFORMATION:

APPLICANT: Itoh, No. 5998170uyuki  
APPLICANT: Martin, Frank  
APPLICANT: Danilenko, Dmitry  
TITLE OF INVENTION: A FIBROBLAST GROWTH FACTOR  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943,915  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mazza, Richard J.

REGISTRATION NUMBER: 27,657  
REFERENCE/DOCKET NUMBER: A-469  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 805.447.4112  
TELEFAX: 805.447.1090  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-943-913-6

Query Match 26.8%; Score 76; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00063;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPPDPLEPR 19  
Db 1 GAPVPPDPLEPR 13

RESULT 12  
US-08-881-037-112  
Sequence 112, Application US/08881037  
Patent No. 6080588  
GENERAL INFORMATION:  
APPLICANT: Glick, Gary D.  
APPLICANT: Swanson, Patrick C.  
TITLE OF INVENTION: DNA BINDING ANTIBODIES  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/881.037  
FILING DATE: 23-JUN-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/443,540  
FILING DATE: 18-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Konksi, Antoinette F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: 203442110710  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 813-5600  
TELEFAX: (650) 494-0792  
TELEX:  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-881-037-112

Query Match 26.8%; Score 76; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00063;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPPDPLEPR 19

Db 1 GAPVPPDPLEPR 13

RESULT 13  
US-08-652-816A-43  
Sequence 43, Application US/08652816A  
Patent No. 5872215  
GENERAL INFORMATION:  
APPLICANT: Osbourn, JK  
APPLICANT: Allen, DJ  
APPLICANT: McCaferly, JG  
TITLE OF INVENTION: Specific binding members, materials and  
METHODS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,816A  
FILING DATE: 23-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125579.4  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125579.8  
FILING DATE: 02-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206318.9  
FILING DATE: 24-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206372.6  
FILING DATE: 23-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9525004.9  
FILING DATE: 07-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610824.6  
FILING DATE: 23-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/02240  
FILING DATE: 02-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/244,597  
FILING DATE: 01-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/33308  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEX:  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-652-816A-43

Query Match 26.8%; Score 76; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVPPDPLEPR 19

Db 23 GAPVPYDDPLEPR 35

RESULT 14  
US-09-142-974B-5

; Sequence 5, Application US/09142974B  
; Patent No. 6451995  
; GENERAL INFORMATION:  
; APPLICANT: Cheung, Nai-Kong V.  
; APPLICANT: Larson, Steven M.  
; APPLICANT: Guo, Hong-Fen  
; APPLICANT: Rivlin, Ken  
; APPLICANT: Sedelain, Michel  
; TITLE OF INVENTION: Single Chain Fv Constructs of Anti-Ganglioside GD2  
; FILE REFERENCE: MSK.P-013-USNP  
; CURRENT APPLICATION NUMBER: US/09/142,974B  
; CURRENT FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: PCT/US97/04427  
; PRIOR FILING DATE: 1997-03-20  
; PRIOR APPLICATION NUMBER: 60/013,703  
; PRIOR FILING DATE: 1996-03-20  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: antibody tag  
US-09-142-974B-5

Query Match 23.9%; Score 68; DB 4; Length 13;  
Best Local Similarity 92.3%; Pred. No. 0.0072;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GAPVPYDDPLEPR 19  
Db 1 GAPVPYDDPLEPR 13

RESULT 15  
US-08-677-010-3  
; Sequence 3, Application US/08677010  
; Patent No. 5925805  
; GENERAL INFORMATION:  
; APPLICANT: Ohlrogge, John B.  
; APPLICANT: Roessler, Keith R.  
; APPLICANT: Shorlosh, Basil S.  
; TITLE OF INVENTION: Structure and Expression of an  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
; STREET: P.O. Box 828  
; CITY: Bloomfield Hills  
; STATE: Michigan  
; COUNTRY: U.S.A.  
; ZIP: 48303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/677,010  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, Deann F.  
; REGISTRATION NUMBER: 36683  
; REFERENCE/DOCKET NUMBER: 6550-00002CFA  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (810)641-1600  
; FAX: (810)641-0270  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2254 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-677-010-3

Query Match 22.2%; Score 63; DB 2; Length 2254;  
Best Local Similarity 27.5%; Pred. No. 17;  
Matches 14; Conservative 12; Mismatches 23; Indels 2; Gaps 1;

QY 1 YLYOMLGAPVPYDDPLEPRREVCLELND--CDELADHIGFOEAYRRFGPV 49  
Db 1818 YIPAYVGPLPVLAPLDPRIIVEYVPENSDDPRAALAGVADNTGKWLGGI 1868

Search completed: December 4, 2002, 15:34:55  
Job time: 15 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:38:50 ; Search time 144 Seconds  
(without alignments)  
219.388 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 49  
Sequence: 1 YLYOMGAVRPPDLEPRR.....DELADHIGQEARRRYGPV 49

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 4569144 seqs, 644733110 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCOTUS\_COMB.pep:\*

2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep:\*

6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep:\*

7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep:\*

8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep:\*

9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep:\*

10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*

11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep:\*

12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep:\*

13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep:\*

14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep:\*

15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep:\*

16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep:\*

17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep:\*

18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep:\*

19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep:\*

20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep:\*

21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep:\*

22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep:\*

23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep:\*

24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep:\*

25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep:\*

26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep:\*

27: /cgn2\_6/ptodata/1/paa/US060\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	49	1	PCT-US02-22821-108
2	49	100.0	49	18	US-09-462-931-2
3	49	100.0	49	25	US-10-157-954-108
4	49	100.0	98	3	US-07-717-811A-5
5	49	100.0	98	6	US-08-246-626-5
6	49	100.0	100	1	PCT-US01-08655-186

7	49	100.0	100	1	PCT-US01-12010-3	Sequence 3, Appl1
8	49	100.0	127	25	US-10-143-899-23	Sequence 23, Appl
9	49	100.0	127	26	US-10-217-651-336	Sequence 336, App
10	42	85.7	42	14	US-09-036-085-5	Sequence 5, Appl1
11	40	81.6	42	27	US-60-160-203-3572	Sequence 3372, Ap
12	40	81.6	42	27	US-60-163-123-1219	Sequence 1219, Ap
13	40	81.6	42	27	US-60-169-840-5089	Sequence 5089, Ap
14	28	57.1	49	13	US-08-973-667-1	Sequence 1, Appl1
15	28	57.1	49	13	US-08-973-667-2	Sequence 2, Appl1
16	28	57.1	49	13	US-08-973-667-3	Sequence 3, Appl1
17	25	51.0	46	20	US-09-657-276-1095	Sequence 1095, Ap
18	25	51.0	47	20	US-09-657-276-1093	Sequence 1093, Ap
19	25	51.0	49	13	US-08-973-667-4	Sequence 4, Appl1
20	22	44.9	140	1	PCT-US01-08655-478	Sequence 478, App
21	20	40.8	21	3	US-07-717-811A-10	Sequence 10, Appl
22	20	40.8	21	6	US-08-246-626-10	Sequence 10, Appl
23	14	28.6	15	3	US-07-717-811A-12	Sequence 12, Appl
24	14	28.6	15	6	US-08-246-626-12	Sequence 12, Appl
25	14	28.6	74	1	PCT-US02-30312-2134	Sequence 2134, Ap
26	14	28.6	74	1	PCT-US02-30412-2134	Sequence 2134, Ap
27	14	28.6	74	23	US-09-962-756-2134	Sequence 2134, Ap
28	14	28.6	74	26	US-10-253-471-2182	Sequence 2182, Ap
29	14	28.6	74	26	US-10-253-493-2134	Sequence 2134, Ap
30	14	28.6	76	1	PCT-US02-30312-2182	Sequence 2182, Ap
31	14	28.6	76	1	PCT-US02-30412-2182	Sequence 2182, Ap
32	14	28.6	76	23	US-09-962-756-2182	Sequence 2182, Ap
33	14	28.6	76	26	US-10-253-471-2182	Sequence 2182, Ap
34	14	28.6	76	26	US-10-253-493-2182	Sequence 2182, Ap
35	13	26.5	13	1	PCT-US00-19843-14	Sequence 14, Appl
36	13	26.5	13	1	PCT-US02-30312-2205	Sequence 2205, Ap
37	13	26.5	13	1	PCT-US02-30412-2205	Sequence 2205, Ap
38	13	26.5	13	1	PCT-US98-17919-6	Sequence 6, Appl1
39	13	26.5	13	12	US-08-844-462-13	Sequence 13, Appl
40	13	26.5	13	14	US-09-011-553-9	Sequence 9, Appl1
41	13	26.5	13	17	US-09-341-550-44	Sequence 44, Appl
42	13	26.5	13	19	US-09-536-556-9	Sequence 9, Appl1
43	13	26.5	13	19	US-09-540-118-29	Sequence 29, Appl
44	13	26.5	13	20	US-09-657-276-1096	Sequence 1096, Ap
45	13	26.5	13	22	US-09-801-968-21	Sequence 21, Appl

#### ALIGNMENTS

RESULT 1  
PCT-US02-22821-108  
Sequence 108, Application PC/TUS0222821  
GENERAL INFORMATION:  
APPLICANT: HK Pharmaceuticals, Inc.  
APPLICANT: Koister, Hubert  
APPLICANT: Siddiqui, Suhail  
TITLE OF INVENTION: Little, Daniel  
TITLE OF INVENTION: Capture Compounds, Collections Thereof  
TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex  
FILE REFERENCE: 24743-2305  
CURRENT APPLICATION NUMBER: PCT/US02/22821  
PRIOR FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: 60/306,019  
PRIOR FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/314,123  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: 60/363,433  
PRIOR FILING DATE: 2002-03-11  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 108  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo Sapien  
PCT-US02-22821-108  
Query Match 100.0%; Score 49; DB 1; Length 49;

Best Local Similarity 100.0%; Pred. No. 1.1e-44;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWIGAVPPYPPDLEPRREVCENLPPDCDELADHIGFOEAYRRFYGPV 49  
DB 1 YLYQWIGAVPPYPPDLEPRREVCENLPPDCDELADHIGFOEAYRRFYGPV 49

## RESULT 2

US-09-462-931-2  
; Sequence 2, Application US/09462931  
; GENERAL INFORMATION:  
; APPLICANT: HELLMAN, Jukka  
; APPLICANT: KITTEN, Sanna-Maria  
; APPLICANT: KARP, Matti  
; APPLICANT: LTVGREN, Timo  
; APPLICANT: VNNEN, Kaleervo  
; APPLICANT: PETERSSON, Kim  
; TITLE OF INVENTION: Isolated osteocalcin fragments  
; FILE REFERENCE: Isolated osteocalcin fragments  
; CURRENT APPLICATION NUMBER: US/09/462,931  
; CURRENT FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: PCT/FI98/00550  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: FI 973371  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 49  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(49)  
; OTHER INFORMATION: Glu at residues 17, 21 and 24 may be gamma-carboxy-Glu  
; US-09-462-931-2

*applied*

Query Match 100.0%; Score 49; DB 18; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.1e-44;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWIGAVPPYPPDLEPRREVCENLPPDCDELADHIGFOEAYRRFYGPV 49  
DB 1 YLYQWIGAVPPYPPDLEPRREVCENLPPDCDELADHIGFOEAYRRFYGPV 49

## RESULT 3

US-10-197-954-108  
; Sequence 108, Application US/10197954  
; GENERAL INFORMATION:  
; APPLICANT: Kister, Hubert  
; APPLICANT: Siddiqi, Subahb  
; APPLICANT: Little, Daniel  
; TITLE OF INVENTION: Capture Compounds, Collections Thereof  
; TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex  
; FILE REFERENCE: 24743-2305  
; CURRENT APPLICATION NUMBER: US/10/197,954  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: 60/306,019  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/314,123  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/363,433  
; PRIOR FILING DATE: 2002-03-11  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 108  
; LENGTH: 49  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
; US-10-197-954-108

Query Match 100.0%; Score 49; DB 25; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.1e-44;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWIGAVPPYPPDLEPRREVCENLPPDCDELADHIGFOEAYRRFYGPV 49  
DB 1 YLYQWIGAVPPYPPDLEPRREVCENLPPDCDELADHIGFOEAYRRFYGPV 49

## RESULT 4

US-07-717-811A-5  
; Sequence 5, Application US/07717811A  
; GENERAL INFORMATION:  
; APPLICANT: Hiroshi EGUCHI et al.  
; TITLE OF INVENTION: Recombinant Human Osteocalcin  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/717,811A  
; FILING DATE: 19910619  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 98 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLER:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY:



LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-717-811A-5

Query Match 100.0%; Score 49; DB 3; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.1e-44;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVYPDPLEPRREVCELNPDCDELADHIGFOEAYRRFGPV 49  
50 YLYQWLGAPVYPDPLEPRREVCELNPDCDELADHIGFOEAYRRFGPV 98

RESULT 5  
US-08-246-626-5  
Sequence 5, Application US/08246626  
GENERAL INFORMATION:  
APPLICANT: Hiroshi EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
City: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 Inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,626  
FILING DATE: 20-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/717,811  
FILING DATE: 19-Jun-1991  
APPLICATION NUMBER: US 08/131,932  
FILING DATE: 05-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 98 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:

ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-246-626-5

Query Match 100.0%; Score 49; DB 6; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.1e-44;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAPVYPDPLEPRREVCELNPDCDELADHIGFOEAYRRFGPV 49  
50 YLYQWLGAPVYPDPLEPRREVCELNPDCDELADHIGFOEAYRRFGPV 98

RESULT 6  
PCT-US01-08655-186  
Sequence 186, Application PC/TUS0108655  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 21272-065  
CURRENT APPLICATION NUMBER: PCT/US01/08655  
CURRENT FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: 09/522,929  
PRIOR FILING DATE: 2000-04-18  
PRIOR APPLICATION NUMBER: 09/770,160  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 09/668,317  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 09/695,783  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 09/728,628  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: 09/783,066  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 09/816,828  
PRIOR FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 584  
SOFTWARE: Custom  
SEQ ID NO 186  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Homo sapiens

PCT-US01-08655-186

Query Match 100.0%; Score 49; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.1e-44;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWIGAPVPYPPDLEPRREVCELNPDDELADHIGFOEAYRRFGPV 49

DB 52 YLYQWIGAPVPYPPDLEPRREVCELNPDDELADHIGFOEAYRRFGPV 100

RESULT 7  
PCT-US01-12010-3

; Sequence 3, Application PC/TUS0112010  
; GENERAL INFORMATION:  
; APPLICANT: Genallsega Pharmaceuticals, Inc.  
; APPLICANT: Benlivesna, Steven C.  
; APPLICANT: Chew, Anne  
; APPLICANT: Choi, Julie Y.  
; APPLICANT: Koshiy, Beena  
; APPLICANT: Stephens, J. Claiborne  
; TITLE OF INVENTION: Haplotypes of the BGLAP Gene  
; FILE REFERENCE: MMH-0512PCT BGLAP  
; CURRENT APPLICATION NUMBER: PCT/US01/12010  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 60/195,840  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Homo sapien  
PCT-US01-12010-3

Query Match 100.0%; Score 49; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.1e-44;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWIGAPVPYPPDLEPRREVCELNPDDELADHIGFOEAYRRFGPV 49

DB 52 YLYQWIGAPVPYPPDLEPRREVCELNPDDELADHIGFOEAYRRFGPV 100

RESULT 8  
US-10-143-899-23

; Sequence 23, Application US/10143899  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT247C1N  
; CURRENT APPLICATION NUMBER: US/10/143,899  
; CURRENT FILING DATE: 2002-05-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (18)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (64)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-143-899-23

Query Match 100.0%; Score 49; DB 25; Length 127;  
Best Local Similarity 100.0%; Pred. No. 2.6e-44;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWIGAPVPYPPDLEPRREVCELNPDDELADHIGFOEAYRRFGPV 49

DB 79 YLYQWIGAPVPYPPDLEPRREVCELNPDDELADHIGFOEAYRRFGPV 127

RESULT 9  
US-10-217-651-336

; Sequence 336, Application US/10217651  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P211C1N  
; CURRENT APPLICATION NUMBER: US/10/217,651  
; CURRENT FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: 60/760,491  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785

;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/244,617  
;; PRIOR FILING DATE: 2000-11-01  
;; PRIOR APPLICATION NUMBER: 60/225,268  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/236,368  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/251,856  
;; PRIOR FILING DATE: 2000-12-08  
;; PRIOR APPLICATION NUMBER: 60/251,868  
;; PRIOR FILING DATE: 2000-12-08  
;; PRIOR APPLICATION NUMBER: 60/229,344  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/234,997  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: 60/229,343  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,345  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,287  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,513  
;; PRIOR FILING DATE: 2000-09-05  
;; PRIOR APPLICATION NUMBER: 60/231,413  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/229,509  
;; PRIOR FILING DATE: 2000-09-05  
;; PRIOR APPLICATION NUMBER: 60/236,367  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/237,039  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,038  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/236,370  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/236,802  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,037  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,040  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/240,960  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/239,935  
;; PRIOR FILING DATE: 2000-10-13  
;; PRIOR APPLICATION NUMBER: 60/239,937  
;; PRIOR FILING DATE: 2000-10-13  
;; PRIOR APPLICATION NUMBER: 60/241,787  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/246,474  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/246,532  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/249,216  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,210  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/226,681  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/225,759  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/225,213  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/227,182  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/225,214  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/235,836  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: 60/230,438  
;; PRIOR FILING DATE: 2000-09-06  
;; PRIOR APPLICATION NUMBER: 60/215,135  
;; PRIOR FILING DATE: 2000-06-30

;; PRIOR APPLICATION NUMBER: 60/225,266  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/249,218  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,208  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,213  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,212  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,207  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,245  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,244  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,217  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,211  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,215  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,264  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,214  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,297  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/232,400  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/231,242  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/232,081  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/232,080  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/231,414  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/231,244  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/233,064  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/233,063  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,397  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,399  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,401  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/241,808  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,826  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,786  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,221  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/246,475  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/231,243  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/233,065

Query Match 100.0%; Score 49; DB 26; Length 127;  
Best Local Similarity 100.0%; Pred. No. 2,6e+44;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVYRDPLEPRREVCELNPDCELDHIGFQEA YRRFGPV 49  
Db 79 YLYQWLGAPVYRDPLEPRREVCELNPDCELDHIGFQEA YRRFGPV 127

```
RESULT 10
US-09-036-085-5
; Sequence 5, Application US/09036085
; GENERAL INFORMATION:
; APPLICANT: GARY S. STEIN ET AL.
; TITLE OF INVENTION: Gene Therapy Using Bone Marrow Transplants Transfected With T
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,085
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PROVISIONAL APPLICATION SERIAL NO. 60/039,839
; FILING DATE: March 6, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DMM-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-085-5

Query Match      85.7%; Score 42; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 3,1e-37;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 APVPYDPLEPRREVCENLPDCCDELADHIGFOEAYRRFYG 49
Db      1 APVPYDPLEPRREVCENLPDCCDELADHIGFOEAYRRFYG 42

RESULT 11
US-60-160-203-3372
; Sequence 3372, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: C1000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3372
; LENGTH: 42
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-3372

Query Match      81.6%; Score 40; DB 27; Length 42;
Best Local Similarity 100.0%; Pred. No. 4.4e-35;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      8 APVPYDPLEPRREVCENLPDCCDELADHIGFOEAYRRFYG 47
Db      3 APVPYDPLEPRREVCENLPDCCDELADHIGFOEAYRRFYG 42

RESULT 12
US-60-163-123-1219
; Sequence 1219, Application US/60163123
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: C1000137
; CURRENT APPLICATION NUMBER: US/60/163,123
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1219
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Human
US-60-163-123-1219

Query Match      81.6%; Score 40; DB 27; Length 42;
Best Local Similarity 100.0%; Pred. No. 4.4e-35;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 APVPYDPLEPRREVCENLPDCCDELADHIGFOEAYRRFYG 47
Db      3 APVPYDPLEPRREVCENLPDCCDELADHIGFOEAYRRFYG 42

RESULT 13
US-60-169-840-5089
; Sequence 5089, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: C1000164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5089
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Human
US-60-169-840-5089

Query Match      81.6%; Score 40; DB 27; Length 42;
Best Local Similarity 100.0%; Pred. No. 4.4e-35;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 APVPYDPLEPRREVCENLPDCCDELADHIGFOEAYRRFYG 47
Db      3 APVPYDPLEPRREVCENLPDCCDELADHIGFOEAYRRFYG 42

RESULT 14
US-08-973-667-1
; Sequence 1, Application US/08973667
; GENERAL INFORMATION:
; APPLICANT: Sakakibara, Shunpei
; APPLICANT: Kimura, Terutoshi
; APPLICANT: Morimoto, Shigeto
; TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 North Glabe Rd. 8th floor
```

```

: CITY: Arlington
: STATE: VA
: COUNTRY: USA
: ZIP: 22201-4741
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/973,667
: FILING DATE: 10-DEC-1997
: CLASSIFICATION: 435
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP97/01246
: FILING DATE: 10-APR-1997
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: JP 8-88608
: FILING DATE: 10-APR-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: JP 9-43331
: FILING DATE: 27-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.
: REGISTRATION NUMBER: 25,327
: REFERENCE/DOCKET NUMBER: 423-43
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-816-4000
: TELEFAX: 703-816-4100
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Human
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 21
: OTHER INFORMATION: /product= "Gla"
:
: US-08-973-667-1
:
: Query Match 57.1%; Score 28; DB 13; Length 49;
: Best Local Similarity 100.0%; Pred. No. 4,2e-22;
: Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 22 VCELNPDCDELADHIGFQEA YRRFYGPV 49
: DB 22 VCELNPDCDELADHIGFQEA YRRFYGPV 49
:
: RESULT 15
: US-08-973-667-2
: GENERAL INFORMATION:
: APPLICANT: Sakakibara, Shunpei
: APPLICANT: Kimura, Terutoshi
: APPLICANT: Morimoto, Shigeto
: TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHAYE P.C.
: STREET: 1100 North Glebe Rd. 8th floor
: CITY: Arlington
: STATE: VA
: COUNTRY: USA
: ZIP: 22201-4741
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/973,667
: FILING DATE: 10-DEC-1997
: CLASSIFICATION: 435
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP97/01246
: FILING DATE: 10-APR-1997
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: JP 8-88608
: FILING DATE: 10-APR-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: JP 9-43331
: FILING DATE: 27-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.
: REGISTRATION NUMBER: 25,327
: REFERENCE/DOCKET NUMBER: 423-43
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-816-4000
: TELEFAX: 703-816-4100
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Human
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 17
: OTHER INFORMATION: /product= "Gla"
:
: US-08-973-667-2
:
: Query Match 57.1%; Score 28; DB 13; Length 49;
: Best Local Similarity 100.0%; Pred. No. 4,2e-22;
: Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 22 VCELNPDCDELADHIGFQEA YRRFYGPV 49
: DB 22 VCELNPDCDELADHIGFQEA YRRFYGPV 49
:
: RESULT 16
: US-08-973-667-3
: GENERAL INFORMATION:
: APPLICANT: Sakakibara, Shunpei
: APPLICANT: Kimura, Terutoshi
: APPLICANT: Morimoto, Shigeto
: TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHAYE P.C.
: STREET: 1100 North Glebe Rd. 8th floor
: CITY: Arlington
: STATE: VA
: COUNTRY: USA
: ZIP: 22201-4741
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/973,667
: FILING DATE: 10-DEC-1997

```

```

;
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JF97/01246
; FILING DATE: 10-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-88608
; FILING DATE: 10-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 9-43331
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 423-43
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 17
; OTHER INFORMATION: /product= "Gla"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 21
; OTHER INFORMATION: /product= "Gla"
;
; US-08-973-667-3
;
Query Match 57.1%; Score 28; DB 13; Length 49;
Best Local Similarity 100.0%; Pred. No. 4,2e-22;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 VCELNPDCDELADHIGFOEAYRRFYGPV 49
DB 22 VCELNPDCDELADHIGFOEAYRRFYGPV 49

RESULT 17
US-09-657-276-1095
; Sequence 1095, Application US/09657276
; GENERAL INFORMATION:
; APPLICANT: Conjuchem, Inc.
; APPLICANT: Bridon, Dominique
; APPLICANT: Ezrin, Alan
; APPLICANT: Milner, Peter
; APPLICANT: Holmes, Darren
; APPLICANT: Thibaudau, Karen
; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
; FILE REFERENCE: 2110
; CURRENT APPLICATION NUMBER: US/09/657,276
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; PRIOR APPLICATION NUMBER: 60/153,406
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 1617
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1095
; LENGTH: 46
; TYPE: PPT
; ORGANISM: Artificial Sequence

```

```

;
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
;
US-09-657-276-1095
;
Query Match 51.0%; Score 25; DB 20; Length 46;
Best Local Similarity 100.0%; Pred. No. 6,7e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFYGPV 49
DB 25 LNPDCDELADHIGFOEAYRRFYGPV 46

RESULT 18
US-09-657-276-1093
; Sequence 1093, Application US/09657276
; GENERAL INFORMATION:
; APPLICANT: Conjuchem, Inc.
; APPLICANT: Bridon, Dominique
; APPLICANT: Ezrin, Alan
; APPLICANT: Milner, Peter
; APPLICANT: Holmes, Darren
; APPLICANT: Thibaudau, Karen
; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
; FILE REFERENCE: 2110
; CURRENT APPLICATION NUMBER: US/09/657,276
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; PRIOR APPLICATION NUMBER: 60/153,406
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 1617
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1093
; LENGTH: 47
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
;
US-09-657-276-1093
;
Query Match 51.0%; Score 25; DB 20; Length 47;
Best Local Similarity 100.0%; Pred. No. 6,9e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFYGPV 49
DB 23 LNPDCDELADHIGFOEAYRRFYGPV 47

RESULT 19
US-08-973-667-4
; Sequence 4, Application US/08973667
; GENERAL INFORMATION:
; APPLICANT: Sakakibara, Shunpei
; APPLICANT: Kimura, Tetsutoshi
; APPLICANT: Morimoto, Shigeto
; TITLE OF INVENTION: ANTI-GLU17-OSTEOCALCIN ANTIBODY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: NIXON & VANDERHAYE P.C.
; STREET: 1100 North Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:

```

```
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/973,667
/ FILING DATE: 10-DEC-1997
/ CLASSIFICATION: 435
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP97/01246
/ FILING DATE: 10-APR-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 8-88608
/ FILING DATE: 10-APR-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 9-43331
/ FILING DATE: 27-FEB-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Crawford, Arthur R.
/ REGISTRATION NUMBER: 25,327
/ REFERENCE/DOCKET NUMBER: 423-43
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-816-4000
/ TELEFAX: 703-816-4100
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 49 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ ORIGINAL SOURCE:
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 17
/ OTHER INFORMATION: /product= "Gla"
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 21
/ OTHER INFORMATION: /product= "Gla"
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 24
/ OTHER INFORMATION: /product= "Gla"
/ US-08-973-667-4

Query Match 51.0%; Score 25; DB 13; Length 49;
Best Local Similarity 100.0%; Pred. No. 7, 2e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LNPDCDELADHIGFOEAYRRFYGPV 49
Db 25 LNPDCDELADHIGFOEAYRRFYGPV 49

RESULT 20
PCT-US01-08655-478
/ Sequence 478, Application PC/TUS0108655
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 21272-065
/ CURRENT APPLICATION NUMBER: PCT/US01/08655
/ CURRENT FILING DATE: 2001-04-16
/ PRIOR APPLICATION NUMBER: 09/522,929
/ PRIOR FILING DATE: 2000-04-18
/ PRIOR APPLICATION NUMBER: 09/770,160
/ PRIOR FILING DATE: 2001-01-26
/ PRIOR APPLICATION NUMBER: 09/668,317
/ PRIOR FILING DATE: 2000-09-22
/ PRIOR APPLICATION NUMBER: 09/695,783
/ PRIOR FILING DATE: 2000-10-24
```

```
/ PRIOR APPLICATION NUMBER: 09/728,628
/ PRIOR FILING DATE: 2000-12-01
/ PRIOR APPLICATION NUMBER: 09/783,066
/ PRIOR FILING DATE: 2001-02-13
/ PRIOR APPLICATION NUMBER: 09/816,828
/ PRIOR FILING DATE: 2001-03-22
/ NUMBER OF SEQ ID NOS: 584
/ SOFTWARE: Custom
/ SEQ ID NO 478
/ LENGTH: 140
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(140)
/ OTHER INFORMATION: Xaa = any amino acid or nothing
/ PCT-US01-08655-478

Query Match 44.9%; Score 22; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 3, 3e-15;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 APVPYDPLEPRREVCELNPDC 29
Db 99 APVPYDPLEPRREVCELNPDC 120

RESULT 21
US-07-717-811A-10
/ Sequence 10, Application US/07717811A
/ GENERAL INFORMATION:
/ APPLICANT: Hiroshi EGUCHI et al.
/ TITLE OF INVENTION: Recombinant Human Osteocalcin
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Wenderoth, Lind & Ponack
/ STREET: 805 Fifteenth Street, N.W., #700
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: DisplayWrite
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/717,811A
/ FILING DATE: 19910619
/ CLASSIFICATION: 435
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren M. Cheek Jr.
/ REGISTRATION NUMBER: 33,367
/ REFERENCE/DOCKET NUMBER:
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-371-8850
/ TELEFAX: 202-371-8856
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 21 amino acids
/ TYPE: AMINO ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE:
/ HYPOTHETICAL:
/ ANTI-SENSE:
/ FRAGMENT TYPE:
/ ORIGINAL SOURCE:
/ ORGANISM:
```

STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-717-811A-10

Query Match 40.8%; Score 20; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 7.6e-14;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 YLYQWLGAPVPPDPLEPRR 20  
Db 1 YLYQWLGAPVPPDPLEPRR 20

RESULT 22  
US-08-246-626-10  
Sequence 10, Application US/08246626  
GENERAL INFORMATION:  
APPLICANT: HIROSHI EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,626  
FILING DATE: 20-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/717,811  
FILING DATE: 19-Jun-1991  
APPLICATION NUMBER: US 08/131,932  
FILING DATE: 05-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.

REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-246-626-10

Query Match 40.8%; Score 20; DB 6; Length 21;  
Best Local Similarity 100.0%; Pred. No. 7.6e-14;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 YLYQWLGAPVPPDPLEPRR 20  
Db 1 YLYQWLGAPVPPDPLEPRR 20

RESULT 23  
US-07-717-811A-12  
Sequence 12, Application US/07717811A  
GENERAL INFORMATION:  
APPLICANT: HIROSHI EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.



COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/717,811A  
FILING DATE: 19910619  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-717-811A-12

Query Match 28.6%; Score 14; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1,6e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 IGFOEAVRRFYGPV 49

Db 2 IGFOEAVRRFYGPV 15  
|||||  
RESULT 24  
US-08-246-626-12  
Sequence 12, Application US/08246626  
GENERAL INFORMATION:  
APPLICANT: HITOSHI EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,626  
FILING DATE: 20-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/717,811  
FILING DATE: 19-Jun-1991  
APPLICATION NUMBER: US 08/131,932  
FILING DATE: 05-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:

```
;
;
;   AUTHORS:
;   TITLE:
;   JOURNAL:
;   VOLUME:
;   ISSUE:
;   PAGES:
;   DATE:
;   DOCUMENT NUMBER:
;   FILING DATE:
;   PUBLICATION DATE:
;   RELEVANT RESIDUES IN SEQ ID NO:
;
US-08-246-626-12

Query Match
Best Local Similarity 100.0%; Score 14; DB 6; Length 15;
Pred. No. 1.6e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 IGFOEAYRRFYGPV 49
    |||
Db 2 IGFOEAYRRFYGPV 15

RESULT 25
PCT-US02-30312-2134
; Sequence 2134, Application PC/TUS0230312
; GENERAL INFORMATION:
; APPLICANT: NOVO NORDISK A/S et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-40578C
; CURRENT APPLICATION NUMBER: PCT/US02/30312
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2134
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
PCT-US02-30312-2134

Query Match
Best Local Similarity 100.0%; Score 14; DB 1; Length 74;
Pred. No. 7.3e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LGAPVYPDPLEPR 19
    |||
Db 61 LGAPVYPDPLEPR 74

RESULT 26
PCT-US02-30412-2134
; Sequence 2134, Application PC/TUS0230412
; GENERAL INFORMATION:
; APPLICANT: NOVO NORDISK A/S et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-40566C
; CURRENT APPLICATION NUMBER: PCT/US02/30412
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2134
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
```

```
PCT-US02-30412-2134

Query Match
Best Local Similarity 100.0%; Score 14; DB 1; Length 74;
Pred. No. 7.3e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LGAPVYPDPLEPR 19
    |||
Db 61 LGAPVYPDPLEPR 74

RESULT 27
US-09-962-756-2134
; Sequence 2134, Application US/09962756
; GENERAL INFORMATION:
; APPLICANT: PILUITLA, RENUKA
; APPLICANT: BRISETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAFER, LADGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-40510S1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2134
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-962-756-2134

Query Match
Best Local Similarity 100.0%; Score 14; DB 23; Length 74;
Pred. No. 7.3e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LGAPVYPDPLEPR 19
    |||
Db 61 LGAPVYPDPLEPR 74

RESULT 28
US-10-253-471-2134
; Sequence 2134, Application US/10253471
; GENERAL INFORMATION:
; APPLICANT: PILUITLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2134
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: peptide  
US-10-253-471-2134

Query Match 28.6%; Score 14; DB 26; Length 74;  
Best Local Similarity 100.0%; Pred. No. 7.3e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVPPDLEPR 19  
DB 61 LGAPVPPDLEPR 74

RESULT 29  
US-10-253-493-2134  
Sequence 2134, Application US/10253493  
GENERAL INFORMATION:  
APPLICANT: PILUTTA, RENUKA et al.  
TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
FILE REFERENCE: 1878-4056  
CURRENT APPLICATION NUMBER: US/10/253,493  
CURRENT FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: 09/962,756  
PRIOR FILING DATE: 2001-09-24  
PRIOR APPLICATION NUMBER: 09/538,038  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 09/146,127  
PRIOR FILING DATE: 1998-09-02  
NUMBER OF SEQ ID NOS: 2227  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 2134  
LENGTH: 74  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-253-493-2134

Query Match 28.6%; Score 14; DB 26; Length 74;  
Best Local Similarity 100.0%; Pred. No. 7.3e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVPPDLEPR 19  
DB 61 LGAPVPPDLEPR 74

RESULT 30  
PCT-US02-30312-2182  
Sequence 2182, Application PC/TUS0230312  
GENERAL INFORMATION:  
APPLICANT: NOVO NORDISK A/S et al.  
TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
FILE REFERENCE: 1878-4057PC  
CURRENT APPLICATION NUMBER: PCT/US02/30312  
CURRENT FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: 09/962,756  
PRIOR FILING DATE: 2001-09-24  
NUMBER OF SEQ ID NOS: 2227  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 2182  
LENGTH: 76  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
PCT-US02-30312-2182

Query Match 28.6%; Score 14; DB 1; Length 76;  
Best Local Similarity 100.0%; Pred. No. 7.5e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVPPDLEPR 19  
DB 63 LGAPVPPDLEPR 76

RESULT 31  
PCT-US02-30412-2182  
Sequence 2182, Application PC/TUS0230412  
GENERAL INFORMATION:  
APPLICANT: NOVO NORDISK A/S et al.  
TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
FILE REFERENCE: 1878-4056PC  
CURRENT APPLICATION NUMBER: PCT/US02/30412  
CURRENT FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: 09/962,756  
PRIOR FILING DATE: 2001-09-24  
NUMBER OF SEQ ID NOS: 2227  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 2182  
LENGTH: 76  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
PCT-US02-30412-2182

Query Match 28.6%; Score 14; DB 1; Length 76;  
Best Local Similarity 100.0%; Pred. No. 7.5e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGAPVPPDLEPR 19  
DB 63 LGAPVPPDLEPR 76

RESULT 32  
US-09-962-756-2182  
Sequence 2182, Application US/09962756  
GENERAL INFORMATION:  
APPLICANT: PILUTTA, RENUKA  
APPLICANT: BRISSETTE, RENEE  
APPLICANT: BLUME, ARTHUR J.  
APPLICANT: SCHAEFER, LAUGE  
APPLICANT: BRANDT, JAKOB  
APPLICANT: GOLDSTEIN, NEIL I.  
APPLICANT: SPETZLER, JANE  
APPLICANT: OSTERGARD, SOREN  
APPLICANT: HANSEN, PER HERTZ  
TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
FILE REFERENCE: 1878-4051US1  
CURRENT APPLICATION NUMBER: US/09/962,756  
CURRENT FILING DATE: 2001-09-24  
PRIOR APPLICATION NUMBER: 09/538,038  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 09/146,127  
PRIOR FILING DATE: 1998-09-02  
NUMBER OF SEQ ID NOS: 2227  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 2182  
LENGTH: 76  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-962-756-2182

Query Match 28.6%; Score 14; DB 23; Length 76;  
Best Local Similarity 100.0%; Pred. No. 7.5e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVYPDPLEPR 19  
|||||  
DB 63 LGAPVYPDPLEPR 76

RESULT 33  
US-10-253-471-2182  
; Sequence 2182, Application US/10253471  
; GENERAL INFORMATION:  
; APPLICANT: PILUTTA, RENKA et al.  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4057  
; CURRENT APPLICATION NUMBER: US/10/253,471  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: 09/962,756  
; PRIOR FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/538,038  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 09/146,127  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2182  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-253-471-2182

Query Match 28.6%; Score 14; DB 26; Length 76;  
Best Local Similarity 100.0%; Pred. No. 7.5e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVYPDPLEPR 19  
|||||  
DB 63 LGAPVYPDPLEPR 76

RESULT 34  
US-10-253-493-2182  
; Sequence 2182, Application US/10253493  
; GENERAL INFORMATION:  
; APPLICANT: PILUTTA, RENKA et al.  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4056  
; CURRENT APPLICATION NUMBER: US/10/253,493  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: 09/962,756  
; PRIOR FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/538,038  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 09/146,127  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2182  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-253-493-2182

Query Match 28.6%; Score 14; DB 26; Length 76;  
Best Local Similarity 100.0%; Pred. No. 7.5e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGAPVYPDPLEPR 19  
|||||

DB 63 LGAPVYPDPLEPR 76

RESULT 35  
PCT-US00-19843-14  
; Sequence 14, Application PC/TUS0019843  
; GENERAL INFORMATION:  
; APPLICANT: Herr, John C.  
; APPLICANT: Norton, Elizabeth J.  
; APPLICANT: Deikman, Alan B.  
; TITLE OF INVENTION: Recombinant Antibody Directed Against Human Sperm  
; TITLE OF INVENTION: Antigen  
; FILE REFERENCE: 00415-02  
; CURRENT APPLICATION NUMBER: PCT/US00/19843  
; CURRENT FILING DATE: 2000-07-21  
; PRIOR APPLICATION NUMBER: 60/145,512  
; PRIOR FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:commercially  
PCT-US00-19843-14

Query Match 26.5%; Score 13; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVYPDPLEPR 19  
|||||  
DB 1 GAPVYPDPLEPR 13

RESULT 36  
PCT-US02-30312-2205  
; Sequence 2205, Application PC/TUS0230312  
; GENERAL INFORMATION:  
; APPLICANT: NOVO NORDISK A/S et al.  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4057PC  
; CURRENT APPLICATION NUMBER: PCT/US02/30312  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: 09/962,756  
; PRIOR FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2205  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
PCT-US02-30312-2205

Query Match 26.5%; Score 13; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GAPVYPDPLEPR 19  
|||||  
DB 1 GAPVYPDPLEPR 13

RESULT 37  
PCT-US02-30412-2205  
; Sequence 2205, Application PC/TUS0230412  
; GENERAL INFORMATION:  
; APPLICANT: NOVO NORDISK A/S et al.

```

; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056PC
; CURRENT APPLICATION NUMBER: PCT/US02/30412
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 2205
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
PCT-US02-30412-2205

Query Match      26.5%; Score 13; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      7 GAVVPYDPDLEPR 19
        |||
Db      1 GAVVPYDPDLEPR 13

RESULT 38
PCT-US98-17919-6
; Sequence 6, Application PC/TUS9817919
; GENERAL INFORMATION:
; APPLICANT: Amgen Inc.
; TITLE OF INVENTION: A FIBROBLAST GROWTH FACTOR
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/17919
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-469
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 805.447.4112
; TELEFAX: 805.447.1090
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US98-17919-6

Query Match      26.5%; Score 13; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      7 GAVVPYDPDLEPR 19
        |||
Db      1 GAVVPYDPDLEPR 13
```

```

RESULT 39
US-08-844-462-13
; Sequence 13, Application US/08844462
; GENERAL INFORMATION:
; APPLICANT: PATTERSON, DALE H.
; APPLICANT: TARR, GEORGE E.
; TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING
; TITLE OF INVENTION: POLYMERS USING MASS SPECTROMETRY.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Patent Administrator - Testa, Hurwitz &
; ADDRESSEE: Thibault
; STREET: High Street Tower, 125 High Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,462
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/796,598
; FILING DATE: 07-FEB-1997
; APPLICATION NUMBER: US 08/446,055
; FILING DATE: 19-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FLYNN Esq., Kerry A.
; REGISTRATION NUMBER: 33,693
; REFERENCE/DOCKET NUMBER: SYP-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-844-462-13

Query Match      26.5%; Score 13; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      7 GAVVPYDPDLEPR 19
        |||
Db      1 GAVVPYDPDLEPR 13

RESULT 40
US-09-011-563-9
; Sequence 9, Application US/09011563
; GENERAL INFORMATION:
; APPLICANT: Wang, Linfa
; TITLE OF INVENTION: EPITOPIC TAGGING SYSTEM
; FILE REFERENCE: Griffith Hack
; CURRENT APPLICATION NUMBER: US/09/011,563
; CURRENT FILING DATE: 1998-06-15
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Description of Artificial Sequence: peptide
; OTHER INFORMATION: sequence
US-09-011-563-9

Query Match          26.5%; Score 13; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 1,7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDLEPR 19
DB 1 GAVPYPPDLEPR 13

RESULT 41
US-09-341-590-44
; Sequence 44, Application US/09341590
; GENERAL INFORMATION:
; APPLICANT: Larsen, Bjarne Due
; TITLE OF INVENTION: PHARMACOLOGICALLY ACTIVE PEPTIDE CONJUGATES HAVING A
; FILE REFERENCE: PPT-20479-US
; CURRENT APPLICATION NUMBER: US/09/341,590
; CURRENT FILING DATE: 1999-07-03
; PRIOR APPLICATION NUMBER: DK 0317/98
; PRIOR FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 44
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: osteocalcin (37-49)
US-09-341-590-44

Query Match          26.5%; Score 13; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 1,7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GFQEAARRRYGPV 49
DB 1 GFQEAARRRYGPV 13

RESULT 42
US-09-536-556-9
; Sequence 9, Application US/09536556
; GENERAL INFORMATION:
; APPLICANT: Wang, Linfa
; TITLE OF INVENTION: EPTIODE TAGGING SYSTEM
; FILE REFERENCE: Griffith Hack
; CURRENT APPLICATION NUMBER: US/09/536,556
; CURRENT FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 09/011,563
; PRIOR FILING DATE: 1998-06-15
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide
US-09-536-556-9

Query Match          26.5%; Score 13; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 1,7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDLEPR 19
DB 1 GAVPYPPDLEPR 13
```

```
RESULT 43
US-09-540-118-29
; Sequence 29, Application US/09540118
; GENERAL INFORMATION:
; APPLICANT: Itoh, Nobuyuki
; TITLE OF INVENTION: Fibroblast Growth Factor-Like Molecules and Uses Thereof
; FILE REFERENCE: 08035,0001-00000
; CURRENT APPLICATION NUMBER: US/09/540,118
; CURRENT FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 29
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:protein tag
US-09-540-118-29

Query Match          26.5%; Score 13; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 1,7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYPPDLEPR 19
DB 1 GAVPYPPDLEPR 13

RESULT 44
US-09-657-276-1096
; Sequence 1096, Application US/09657276
; GENERAL INFORMATION:
; APPLICANT: Conjuchem, Inc.
; APPLICANT: Bridon, Dominique
; APPLICANT: Ezrin, Alan
; APPLICANT: Milner, Peter
; APPLICANT: Holmes, Darren
; APPLICANT: Thibaudau, Karen
; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
; FILE REFERENCE: 2110
; CURRENT APPLICATION NUMBER: US/09/657,276
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/134,406
; PRIOR FILING DATE: 1999-05-17
; PRIOR APPLICATION NUMBER: 60/153,406
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 1617
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1096
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-657-276-1096

Query Match          26.5%; Score 13; DB 20; Length 13;
Best Local Similarity 100.0%; Pred. No. 1,7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GFQEAARRRYGPV 49
DB 1 GFQEAARRRYGPV 13

RESULT 45
```

US-09-801-968-21  
; Sequence 21, Application US/09801968  
; GENERAL INFORMATION:  
; APPLICANT: Itoh, Nobuyuki  
; APPLICANT: Kavanaugh, W. Michael  
; TITLE OF INVENTION: HUMAN FGF-23 GENE AND GENE EXPRESSION  
; TITLE OF INVENTION: PRODUCTS  
; FILE REFERENCE: PP-17150.001/201130.40901  
; CURRENT APPLICATION NUMBER: US/09/801,968  
; CURRENT FILING DATE: 2001-03-07  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: E tag  
US-09-801-968-21

Query Match 26.5%; Score 13; DB 22; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAPVPYDPLEPR 19  
|||||  
Db 1 GAPVPYDPLEPR 13

Search completed: December 4, 2002, 15:42:55  
Job time : 146 secs







PA (KARP/) KARP M.  
 PA (LOEV/) LOEVGREN T.  
 PA (PERT/) PETERSSON K.  
 PA (VAAE/) VAAEVAENEN H K.  
 XX  
 PI Hellman J, Kaekenan S, Karp M, Loevgren T, Pettersson K;  
 PI Vaaenaaenen HK;  
 XX  
 DR WPI: 1999-180972/15.  
 DR N-PSDB: AAX19767.  
 XX  
 PT Human osteocalcin peptide fragments - useful for detecting  
 PT metabolic bone turnover rates and disorders  
 XX  
 PS Claim 1: Fig 1A; 49pp; English.  
 CC An assay has been developed for the detection of gamma-carboxylated  
 CC human osteocalcin found in urine, where the glutamic acid residues at  
 CC positions 17, 21 and 24 are gamma-carboxylated. The assay is used to  
 CC measure the rate of bone turnover (formation and/or resorption) and/or  
 CC for investigating metabolic bone disorders in individuals. Especially  
 CC mentioned, the assays can detect differences between adults and children  
 CC going through puberty, pre- and post-menopausal women and children  
 CC having high bone turnover. The assay allows more sensitive detection of  
 CC human osteocalcin (hOC) in urine than in serum. Significant increases in  
 CC hOC (40-48%) were detected in serum in menopause (normal increase is  
 CC 30-50% above pre-menopausal women). Detection of urine hOC using the  
 CC assay showed an increase as high as 75-79%. The present sequence  
 CC represents a fragment of hOC.  
 XX  
 SQ Sequence 49 AA:

Query Match 100.0%; Score 284; DB 20; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-30;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPPDPLEPRREVCENPPDCDELADHIGFQEAAYRRFGPV 49  
 DB 1 YLYQWLGAPVPPDPLEPRREVCENPPDCDELADHIGFQEAAYRRFGPV 49

RESULT 2  
 AAR10146  
 ID AAR10146 standard; peptide: 51 AA.  
 XX  
 AC AAR10146;  
 XX  
 DT 27-MAR-1991 (first entry)  
 XX  
 DE Human osteocalcin precursor polypeptide (I).  
 XX  
 KW Human osteocalcin precursor polypeptide; OC; carboxypeptidase B;  
 KW calcium; vitamin K; bone formation; dysbolism.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..1 /note= "residue may be omitted"  
 FT Misc-difference 51..51 /note= "residue may be omitted"  
 FT  
 XX  
 PN JP02201294-A.  
 XX  
 PD 03-DEC-1990.  
 XX  
 PF 22-AUG-1989; 89JP-0214239.  
 XX  
 PR 06-DEC-1988; 88JP-0306931.  
 PR 22-AUG-1989; 89JP-0214239.  
 XX  
 PA (TAKA-) TAKARA SYUZO KK.  
 XX

1024b)

DR WPI: 1991-018865/03.  
 XX  
 PT Human osteocalcin (OC) precursor polypeptide - used to prepare  
 PT purified OC by digestion with carboxypeptidase B.  
 XX  
 PS Claim 1: Page 1; 11pp; Japanese.  
 CC  
 CC To K1 is attached H and to K51 is attached OH.  
 CC A novel gene encoding human OC precursor polymer was inserted into  
 CC plasmid pOC 980. E.coli HB101 was transformed with this plasmid, and  
 CC made to efficiently produce the objective polymer. The polymer was  
 CC treated with lysylendopeptidase B to remove spacers. The obtained  
 CC human precursor was further treated with carboxypeptidase to prepare a  
 CC purified human OC.  
 CC OC is a calcium bonded protein (depending on vitamin K) produced in the  
 CC bone. It is thought to be a promoting factor during the bone  
 CC formation and may be used to treat diseases due to dysbolism of the  
 CC bone.  
 CC See also AAR10147 and AAO10193-98.  
 CC  
 SQ Sequence 51 AA:

Query Match 100.0%; Score 284; DB 12; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-30;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YLYQWLGAPVPPDPLEPRREVCENPPDCDELADHIGFQEAAYRRFGPV 49  
 DB 2 YLYQWLGAPVPPDPLEPRREVCENPPDCDELADHIGFQEAAYRRFGPV 50

## RESULT 3

AAW76094  
 ID AAW76094 standard; Protein: 98 AA.

XX  
 AC AAW76094;  
 XX

DT 21-DEC-1998 (first entry)  
 XX

DE Human osteocalcin protein.  
 XX

KW Promoter; tissue-specific gene expression; skeletal tissue;  
 KW stem cell; bone; cartilage; osteocalcin; hOC gene; human;  
 KW osteoporosis; osteopenia; osteosarcoma; cancer; metastasis;  
 KW gene therapy.  
 XX

OS Homo sapiens.  
 XX

PN WO9839427-A2.  
 XX

PD 11-SEP-1998.  
 XX

PF 06-MAR-1998; 98MO-US04421.  
 XX

PR 06-MAR-1997; 97US-0039839.  
 XX

PA (UYMA-) UNIV MASSACHUSETTS.  
 XX

PI Frenkel B, Hou Z, Lian JB, Nilsson S, Peters S;  
 PI Quesenberry P, Stein GS, Stein JL;  
 XX

DR WPI: 1998-495839/42.  
 XX

N-PSDB: AAV46429.  
 XX

PT Expression of exogenous genes in differentiated cells - by  
 PT transducing pluripotent stem cells capable of maturing into  
 PT differentiated cells with nucleic acid comprising exogenous gene,  
 PT useful for, e.g. treatment of osteoporosis  
 XX

PS Disclosure: Page 33-34; 63pp; English.  
 XX

CC This is the amino acid sequence of the human osteocalcin, deduced  
 CC from the coding exons of the human hOC gene (see AAV46429). The

1024a)

CC Invention pertains to a method for expressing endogenous genes in  
 CC differentiated cells of a specific type. The method involves  
 CC contacting pluripotent stem cells capable of maturing into  
 CC differentiated cells with a nucleic acid comprising an exogenous  
 CC gene linked to a regulatory element capable of controlling expression  
 CC of the exogenous gene in the differentiated cells. A population of  
 CC transduced stem cells capable of maturing into differentiated cells  
 CC expressing the exogenous gene is produced. Preferably, the  
 CC differentiated cells are in a tissue of interest, such as bone or  
 CC cartilage, and the exogenous gene is operably linked to at least one  
 CC osteocalcin regulatory element derived from the hOC promoter. The  
 CC exogenous gene can encode a therapeutic protein useful for treating  
 CC a disease, especially osteoporosis, osteopenia, osteosarcoma,  
 CC primary malignancy or metastases (all claimed).

XX Sequence 98 AA;

Query Match 100.0%; Score 284; DB 19; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-30;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPPLPRREVCELNPCDELADHIGFOEAYRRFGPV 49  
 |||  
 Db 50 YLYQWLGAAPVPPPLPRREVCELNPCDELADHIGFOEAYRRFGPV 98

RESULT 4

AA010147 standard; Protein; 100 AA.

XX AAR10147;

DT 27-MAR-1991 (first entry)

XX Human osteocarcin precursor polypeptide (II).

KM Human osteocarcin precursor polypeptide; OC; carboxypeptidase B;

KW calcium; vitamin K; bone formation; dysbolism.

OS Homo sapiens.

PN JF02201294-A.

PD 03-DEC-1990.

PF 22-AUG-1989; 89JP-0214239.

PR 06-DEC-1988; 88JP-0306931.

XX 22-AUG-1989; 89JP-0214239.

PA (TAKA-) TAKARA SYUZO KK.

DR WPI; 1991-01865/03.

XX Human osteocarcin (OC) precursor polypeptide - used to prepare

PT purified OC by digestion with carboxypeptidase B.

PS Claim 5; Page 1; 11pp; Japanese.

XX To K1 is attached H and to V100 is attached OH.

CC A novel gene encoding human OC precursor polymer was inserted into

CC plasmid pOC 980. E.coli HB101 was transformed with this plasmid, and

CC made to efficiently produce the objective polymer. The polymer was

CC treated with lysine endopeptidase B to remove spacers. The obtained

CC human precursor was further treated with carboxy peptidase to prepare a

CC purified human OC.

CC OC is a calcium bonded protein (depending on vitamin K) produced in the

CC bone. It is thought to be a promoting factor during the bone

CC formation and may be used to treat diseases due to dysbolism of the

Query Match 100.0%; Score 284; DB 12; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-30;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPPLPRREVCELNPCDELADHIGFOEAYRRFGPV 49  
 |||  
 Db 52 YLYQWLGAAPVPPPLPRREVCELNPCDELADHIGFOEAYRRFGPV 100

RESULT 5

AAU010687 standard; Protein; 100 AA.

XX AAU010687;

DT 14-FEB-2002 (first entry)

DE Reference sequence for human BGLAP protein.

KW Human; single nucleotide polymorphism; SNP; BGLAP; chromosome 1q25-q31;

KW bone gamma carboxyglutamate protein; haplotyping; genotyping;

XX osteoporosis; osteopathic.

OS Homo sapiens.

PN WO200177131-A2.

PD 18-OCN-2001.

PF 11-APR-2001; 2001WO-US12010.

PR 11-APR-2000; 2000US-195840P.

XX (GENA-) GENAISSANCE PHARM INC.

PA Bentivegna SC, Chew A, Choi JY, Koshy B, Rounds E, Stephens JC;

DR WPI; 2002-041288/05.

XX N-PSDB; AAS16363, AAS16364.

XX New haplotypes of the human bone gamma carboxyglutamate protein gene,

PT useful to diagnose and treat diseases associated with the gene such as

PS osteoporosis

XX Claim 27; Fig 3; 53pp; English.

CC The present invention relates to novel single nucleotide polymorphisms

CC (SNPs) in the human bone gamma carboxyglutamate protein (BGLAP) gene

CC located on chromosome 1q25-q31, and methods for haplotyping and/or

CC genotyping the BGLAP gene in an individual. The methods of the

CC invention make use of allele-specific oligonucleotides (ASOs) as probes

CC and primers and/or primer-extension oligonucleotides for detecting the

CC BGLAP gene polymorphisms. The polymorphisms and screened compounds are

CC useful for (developing) treatment of diseases associated with BGLAP

CC activity, such as osteoporosis. The present sequence represents a

XX reference sequence for the BGLAP protein.

XX Sequence 100 AA;

Query Match 100.0%; Score 284; DB 23; Length 100;

Best Local Similarity 100.0%; Pred. No. 8.3e-30;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPPLPRREVCELNPCDELADHIGFOEAYRRFGPV 49  
 |||  
 Db 52 YLYQWLGAAPVPPPLPRREVCELNPCDELADHIGFOEAYRRFGPV 100

NOT related  
 to 05/05/02  
 do not 12/10/02

AC AAM34263;  
XX 23-APR-1998 (first entry)  
XX Glu17-osteocalcin peptide #1.  
DE Glu17-osteocalcin peptide #1.  
XX  
XX Glu17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
KM osteoporosis; human.  
XX  
XX Homo sapiens.  
OS  
XX  
XX  
FH Key Location/Qualifiers  
FT Modified-site 21 /note="gamma-carboxyglutamic acid"  
FT  
XX  
XX W09738309-A1.  
XX  
XX 16-OCT-1997.  
XX  
XX 10-APR-1997; 97WO-JP01246.  
XX  
XX 27-FEB-1997; 97JP-0043331.  
XX 10-APR-1996; 96JP-0088608.  
XX  
XX (EISA ) EISAI CO LTD.  
XX  
XX Kimura T, Morimoto S, Sakakibara S;  
PI WPI: 1997-512875/47.  
XX  
XX Antibody specific for Glu17-osteocalcin, or its fragment - for  
PT diagnosis of bone disorders such as osteoporosis  
XX  
XX Claim 3; Page 16; 28pp; Japanese.  
PS  
XX This sequence represents a human Glu17-osteocalcin peptide. This sequence  
CC is recognised by the antibody of the invention. The antibody of the  
CC invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
CC binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The  
CC antibody can be used for the diagnosis of bone related disorders, such as  
CC osteoporosis.  
XX  
XX Sequence 49 AA;  
SQ  
Query Match 97.9%; Score 278; DB 18; Length 49;  
Best Local Similarity 98.0%; Pred. No. 2.3e-29;  
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YLYQWLGAVPYPDPLEPRREVCELNPCDELADHIGFQEAAYRRFGPV 49  
DB 1 YLYQWLGAVPYPDPLEPRREVCELNPCDELADHIGFQEAAYRRFGPV 49

RESULT 7  
AAR20043  
ID AAR20043 standard; Protein; 97 AA.  
XX  
XX AAR20043;  
AC  
XX 09-APR-1992 (first entry)  
DT  
XX  
XX Fusion protein for expression of human osteocalcin.  
DE  
XX gamma-carboxyglutamic acid; bone matrix; Gla protein.  
XX  
XX  
OS Synthetic.  
XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..22 /label= signal\_peptide  
FT Peptide 23..44 /label= pro\_peptide  
FT Cleavage-site 45..48

FT Protein 49..97  
FT /label= osteocalcin  
XX  
XX EP463571-A.  
XX  
XX 02-JAN-1992.  
XX  
XX 20-JUN-1991; 91EP-0110173.  
XX  
XX 20-JUN-1991; 91EP-0110173.  
XX  
XX 30-NOV-1990; 90JP-0330146.  
XX 20-JUN-1990; 90JP-0159909.  
XX  
XX (TEIJ ) TEIJIN KK.  
XX  
XX Eguchi H, Kamimura TF, Sugiyama T, Hosoda K;  
PI WPI: 1992-009183/02.  
XX  
XX N-PSDB; AAO20210.  
XX  
XX Human osteocalcin produ. - using DNA coding for human osteocalcin  
PT fusion protein for expression in host cells  
XX  
XX Claim 15; Fig 7; 53pp; English.  
PS  
XX This sequence is a specific example of a claimed generic fusion  
CC protein comprising human osteocalcin. The pro-peptide is recognised  
CC by an enzyme capable of Glu to Gla conversion on human osteocalcin.  
CC The recombinant protein was obtained by culturing host cells  
CC transformed with a vector containing the synthetic coding sequence.  
CC The Glu residues could then be converted to Gla (i.e.  
CC gamma-carboxyglutamic acid) and the osteocalcin sequence cleaved  
CC from the propeptide. The mature protein is suitable for use in  
CC immunoassays and as a drug for treatment of bone metabolism  
CC disorders. See also AAR20044-6.  
XX  
XX Sequence 97 AA;  
SQ  
Query Match 97.5%; Score 277; DB 13; Length 97;  
Best Local Similarity 98.0%; Pred. No. 6.8e-29;  
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YLYQWLGAVPYPDPLEPRREVCELNPCDELADHIGFQEAAYRRFGPV 49  
DB 49 YLYQWLGAVPYPDPLEPRREVCELNPCDELADHIGFQEAAYRRFGPV 97

RESULT 8  
AAU10688  
ID AAU10688 standard; Protein; 100 AA.  
XX  
XX AAU10688;  
AC  
XX 14-FEB-2002 (first entry)  
DT  
XX  
XX Polymorphic variant of human BGLAP protein.  
DE  
XX  
XX Human; single nucleotide polymorphism; SNP; BGLAP; chromosome 1q25-q31;  
XX bone gamma carboxyglutamate protein; haplotyping; genotyping;  
XX osteoporosis; osteopathic; variant.  
XX  
XX Homo sapiens.  
OS  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 82 /note="Substitution of Glu to Lys"  
FT Misc-difference 94 /note="Substitution of Arg to Gln"  
XX  
XX W0200177131-A2.  
XX  
XX 18-OCT-2001.  
XX  
XX 11-APR-2001; 2001WO-US12010.

XX	11-APR-2000; 2000US-195840P.	
PR	(GENA-) GENAISSANCE PHARM INC.	
PA	Bentivegna SC, Chew A, Choi JY, Koshy B, Rounds E, Stephens JC;	
XX	WPI: 2002-041288/05.	
DR		
XX	New haplotypes of the human bone gamma carboxyglutamate protein gene,	
PT	useful to diagnose and treat diseases associated with the gene such as	
PR	osteoporosis	
PS		
XX	Claim 27; Page -: 53pp; English.	
XX		
CC	The present invention relates to novel single nucleotide polymorphisms	
CC	(SNPs) in the human bone gamma carboxyglutamate protein (BGLAP) gene	
CC	located on chromosome 1q25-q31, and methods for haplotyping and/or	
CC	genotyping the BGLAP gene in an individual. The methods of the	
CC	invention make use of allele-specific oligonucleotides (ASOs) as probes	
CC	and primers and/or primer-extension oligonucleotides for detecting the	
CC	BGLAP gene polymorphisms. The polynucleotides and screened compounds are	
CC	useful for (developing) treatment of diseases associated with BGLAP	
CC	activity, such as osteoporosis. The present sequence represents a	
CC	polymorphic variant of the BGLAP protein (AAU10687).	
CC	Note: The present sequence is not given in the specification but is	
CC	created by the indexer from the information given in the patent.	
SO	Sequence 100 AA;	
Query Match	97.2%; Score 276; DB 23; Length 100;	
Best Local Similarity	95.9%; Fred. No. 9.6e-29;	
Matches 47; Conservative 2; Mismatches 0; Indels 0; Gaps 0;		
OY	1 YLYOMIGAPVPYPDPLEPREVCEINPDCDELADHIGFOEAYRRFGPV 49	
DB	52 YLYOMIGAPVPYPDPLEPREVCEINPDCDKLADHIGFOEAYRRFGPV 100	
RESULT 9		
AAK32936		
ID	AAK32936 standard; peptide: 49 AA.	
XX		
AC	AAK32936:	
DT	05-JUL-1993 (first entry)	
XX		
DE	21, 24, G1a human osteocalcin peptide.	
XX		
KW	Gamma-carboxyglutamic acid; OS.	
XX		
OS	Synthetic.	
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 21	
FT	/label= OTHER	
FT	/note= "OTHER= gamma-carboxyglutamic acid"	
FT	Misc-difference 24	
FT	/label= OTHER	
FT	/note= "OTHER= gamma-carboxyglutamic acid"	
XX		
PN	JP05032697-A.	
XX		
PD	09-FEB-1993.	
XX		
PE	31-JUL-1991; 91JP-0213251.	
XX		
PR	31-JUL-1991; 91JP-0213251.	
XX		
PA	(TEIJ ) TEIJIN LTD.	
XX		
DR	WPI: 1993-088665/11.	
XX		

[illegible]

Query Match 95.8%; Score 272; DB 18; Length 49;  
Best Local Similarity 95.9%; Pred. No. 1.4e-28;  
Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YLYQWLGAIPVPPDPLEPRREYCELNPDDELADHIGFOEAYRRFGPV 49  
|||||  
DB 1 YLYQWLGAIPVPPDPLEPRRXVCXLPDDELADHIGFOEAYRRFGPV 49

RESULT 11  
AAM34265  
ID AAM34265 standard; peptide: 49 AA.  
AC AAM34265;  
XX  
DT 23-APR-1998 (first entry)  
XX  
DE Gla21-osteocalcin peptide #1.  
XX  
KW Gla17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
KM osteoporosis; human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 17  
FT Modified-site /note= "gamma-carboxyglutamic acid"  
FT Modified-site 21  
FT Modified-site /note= "gamma-carboxyglutamic acid"  
XX  
FN WO9738309-A1.  
XX  
PD 16-OCT-1997.  
XX  
PF 10-APR-1997; 97WO-JP01246.  
XX  
PR 27-FEB-1997; 97JP-0043331.  
XX  
PR 10-APR-1996; 96JP-0088608.  
XX  
PA (EISA ) EISAI CO LTD.  
XX  
PI Kimura T, Morimoto S, Sakakibara S;  
XX  
DR WPI: 1997-512875/47.  
XX  
PT Antibody specific for Gla17-osteocalcin, or its fragment - for  
PT diagnosis of bone disorders such as osteoporosis  
XX  
PS Claim 4; Page 17; 28pp; Japanese.  
XX  
CC This sequence represents a human Gla21-osteocalcin peptide. This sequence  
CC is recognised by the antibody of the invention. The antibody of the  
CC invention is an anti-Gla17-osteocalcin antibody or its fragment, which  
CC binds to Gla17-osteocalcin, Gla21-osteocalcin or their fragments. The  
CC antibody can be used for the diagnosis of bone related disorders, such as  
CC osteoporosis.  
XX  
SQ Sequence 49 AA;

Query Match 95.8%; Score 272; DB 18; Length 49;  
Best Local Similarity 95.9%; Pred. No. 1.4e-28;  
Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YLYQWLGAIPVPPDPLEPRREYCELNPDDELADHIGFOEAYRRFGPV 49  
|||||  
DB 1 YLYQWLGAIPVPPDPLEPRRXVCXLPDDELADHIGFOEAYRRFGPV 49

RESULT 12  
AAR32937  
ID AAR32937 standard; peptide: 49 AA.  
XX  
AC AAR32937;

XX 05-JUL-1993 (first entry)  
DT 17, 21, 24, Gla human osteocalcin peptide.  
XX  
DE Gamma-carboxyglutamic acid; OS.  
XX  
KM Synthetic.  
XX  
OS  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 17  
FT Misc-difference /label= OTHER  
FT Misc-difference 21  
FT Misc-difference /note= "OTHER- gamma-carboxyglutamic acid"  
FT Misc-difference 24  
FT Misc-difference /note= "OTHER- gamma-carboxyglutamic acid"  
FT Misc-difference 24  
FT Misc-difference /label= OTHER  
FT Misc-difference /note= "OTHER- gamma-carboxyglutamic acid"  
XX  
PN JP05032697-A.  
XX  
PD 09-FEB-1993.  
XX  
PF 31-JUL-1991; 91JP-0213251.  
XX  
PR 31-JUL-1991; 91JP-0213251.  
XX  
PA (TEIJIN ) TEIJIN LTD.  
XX  
DR WPI: 1993-088665/11.  
XX  
PT Synthetic human osteocalcin for standard in determ. of natural  
PT osteocalcin - prep. by introducing gamma-carboxy:glutamic acid  
PT as fluorenyl protected gp.  
XX  
PS Claim 2; Page 2; 10pp; Japanese.  
XX  
CC The synthetic 17, 21, 24, gamma-carboxyglutamic acid form of human  
CC osteocalcin (OS) was produced by introducing protected Gla. The  
CC substance may be produced in high yield and is useful as standard  
CC for the determination of human OS.  
XX  
SQ Sequence 49 AA;

Query Match 93.7%; Score 266; DB 14; Length 49;  
Best Local Similarity 93.9%; Pred. No. 8.8e-28;  
Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 YLYQWLGAIPVPPDPLEPRREYCELNPDDELADHIGFOEAYRRFGPV 49  
|||||  
DB 1 YLYQWLGAIPVPPDPLEPRRXVCXLPDDELADHIGFOEAYRRFGPV 49

RESULT 13  
AAM34266  
ID AAM34266 standard; peptide: 49 AA.  
XX  
AC AAM34266;  
XX  
DT 23-APR-1998 (first entry)  
XX  
DE Gla21-osteocalcin peptide #2.  
XX  
KW Gla17-osteocalcin; Gla21-osteocalcin; antibody; bone disorder; diagnosis;  
KM osteoporosis; human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 17  
FT Modified-site /note= "gamma-carboxyglutamic acid"

FT Modified-site 21 /note- "gamma-carboxyglutamic acid"  
 FT Modified-site 24 /note- "gamma-carboxyglutamic acid"  
 FT  
 XX MO9738309-A1.  
 XX  
 PD 16-OCT-1997.  
 XX  
 PF 10-APR-1997; 97WO-JP01246.  
 XX  
 PR 27-FEB-1997; 97JP-0043331.  
 PR 10-APR-1996; 96JP-0088608.  
 XX  
 PA (EISA ) EISAI CO LTD.  
 XX  
 PI Kimura T, Morimoto S, Sakakibara S;  
 DR WPI; 1997-512875/47.  
 XX  
 PT Antibody specific for Glu17-osteocalcin, or its fragment - for  
 PT diagnosis of bone disorders such as osteoporosis  
 XX  
 PS Claim 4; Page 18; 28pp; Japanese.  
 XX  
 CC This sequence represents a human Gla21-osteocalcin peptide. This sequence  
 CC is recognised by the antibody of the invention. The antibody of the  
 CC invention is an anti-Glu17-osteocalcin antibody or its fragment, which  
 CC binds to Glu17-osteocalcin, Gla21-osteocalcin or their fragments. The  
 CC antibody can be used for the diagnosis of bone related disorders, such as  
 CC osteoporosis.  
 XX  
 SQ Sequence 49 AA;

Query Match 93.7%; Score 266; DB 18; Length 49;  
 Best Local Similarity 93.9%; Pred. No. 8.8e-28;  
 Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPDLEPRREVCELNPDCDELADHIGFOEAYRRFGPV 49  
 |||||  
 Db 1 YLYQWLGAAPVPPDLEPRRXVCXKLNPCDELADHIGFOEAYRRFGPV 49

## RESULT 14

AAW01681  
 ID AAW01681 standard; protein; 49 AA.

AC AAW01681;

DT 01-APR-1997 (first entry)

DE Bone Gla protein.

XX BGP; bone gla protein; osteocalcin; Vitamin K-dependent protein;  
 KW bone matrix; therapy; diagnosis; assay; metabolic bone disease.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers

FT MISC-difference 1 /note- "underlined in specification"

FT MISC-difference 3 /note- "underlined in specification"

FT MISC-difference 12 /note- "underlined in specification"

FT Modified-site 21 /label- OTHER

FT MISC-difference 23..29 /note- "gamma-carboxyglutamic acid"

FT Modified-site 24 /note- "underlined in specification"

FT /label- OTHER  
 /note- "gamma-carboxyglutamic acid"

FT MISC-difference 42 /note- "underlined in specification"  
 FT MISC-difference 46 /note- "underlined in specification"  
 FT  
 XX DE4340597-A1.  
 XX  
 PD 01-JUN-1995.  
 XX  
 PF 29-NOV-1993; 93DE-4340597.  
 XX  
 PR 29-NOV-1993; 93DE-4340597.  
 XX  
 PA (HENN-) HENNING BERLIN GMBH.  
 PA (BRAH-) BRAHMS DIAGNOSTICA GMBH.  
 XX  
 PI Bergmann A, Weckermann R;  
 DR WPI; 1995-201516/27.  
 XX  
 PT Determ. of osteocalcin in serum or plasma - with addn. of divalent  
 PT metal ions to inhibit decompn. of the protein, useful in therapy  
 PT and diagnosis of bone disease  
 XX  
 PS Disclosure; Column 1; 9pp; German.

XX A method for determining concentration of osteocalcin in serum or plasma  
 CC is improved with addition of divalent metal ions to inhibit decomposition  
 CC of the protein. The method is useful in therapy and diagnosis of bone  
 CC disease. The present sequence is osteocalcin (a vitamin K-dependent  
 CC protein) that is a component of the bone matrix, or alternatively bone Gla  
 CC protein).  
 XX  
 SQ Sequence 49 AA;

Query Match 93.0%; Score 264; DB 16; Length 49;  
 Best Local Similarity 93.9%; Pred. No. 1.6e-27;  
 Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 YLYQWLGAAPVPPDLEPRREVCELNPDCDELADHIGFOEAYRRFGPV 49  
 |||||  
 Db 1 YLYQWLGAAPVPPDLEPRRXVCXKLNPCDELADHIGFOEAYRRFGPV 49

## RESULT 15

AAB91917  
 ID AAB91917 standard; Peptide; 47 AA.

AC AAB91917;

DT 22-JUN-2001 (first entry)

DE Bone Gla protein peptide SEQ ID NO:1093.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidy1; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 XX  
 OS Homo sapiens.

OS Synthetic.

PN WO200069900-A2.

PN 23-NOV-2000.

PF 17-MAY-2000; 2000WO-US13576.

PR 17-MAY-1999; 99US-0134406.

PR 10-SEP-1999; 99US-0153406.

PR 15-OCT-1999; 99US-0159783.

PA (CONU-) CONJUCHEM INC.

PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
 XX  
 DR WPI; 2001-112059/12.

XX  
 PT Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity

XX  
 PS Disclosure; Page 552-553; 733pp; English.

XX  
 CC The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimido and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.

XX  
 SQ Sequence 47 AA;

Query Match 89.1%; Score 253; DB 22; Length 47;

Best Local Similarity 95.9%; Pred. No. 4.4e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 1 YLQWIGAPVPYPPDPEPRREVCLELNDDELADHIGFOEAYRREGPV 49  
 |||||||  
 DB 1 YLQWIGAPVPYPPDPEPRR-VC-LNPDCDELADHIGFOEAYRREGPV 47

Search completed: December 4, 2002, 15:33:14  
 Job time : 35 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:33:19 ; Search time 141 Seconds  
(without alignments)

224.056 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 284  
Sequence: 1 YLXQMLGAVPPYDPLEPRR.....DELADHIGFOEATRRPYGPV 49

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 4569144 segs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US08.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	284	100.0	49	PCT-US02-22821-108
2	284	100.0	49	US-09-462-931-2
3	284	100.0	49	US-10-197-954-108
4	284	100.0	98	US-07-717-811A-5
5	284	100.0	98	US-08-246-626-5
6	284	100.0	100	PCT-US01-08655-166

7	284	100.0	100	PCT-US01-12010-3	Sequence 3, Appl1
8	284	100.0	127	US-10-143-899-23	Sequence 23, Appl1
9	284	100.0	127	US-10-217-651-336	Sequence 336, Appl1
10	278	97.9	49	US-08-973-667-1	Sequence 1, Appl1
11	272	95.8	49	US-08-973-667-2	Sequence 2, Appl1
12	272	95.8	49	US-08-973-667-3	Sequence 3, Appl1
13	266	93.7	49	US-08-973-667-4	Sequence 4, Appl1
14	253	89.1	47	US-09-657-276-1093	Sequence 1093, Ap
15	240	84.5	42	US-09-036-085-5	Sequence 5, Appl1
16	237.5	83.6	46	US-09-657-276-1095	Sequence 1095, Ap
17	237	83.5	140	PCT-US01-08655-478	Sequence 478, App
18	231	81.3	42	US-60-160-203-3372	Sequence 3372, Ap
19	231	81.3	42	US-60-163-123-1219	Sequence 1219, Ap
20	231	81.3	42	US-60-169-840-5089	Sequence 5089, Ap
21	203	71.5	43	US-09-036-085-22	Sequence 22, Appl
22	119	41.9	21	US-07-717-811A-10	Sequence 10, Appl
23	119	41.9	21	US-08-246-626-10	Sequence 10, Appl
24	86	30.3	20	US-09-036-085-10	Sequence 10, Appl
25	84	29.6	20	US-09-036-085-16	Sequence 16, Appl
26	81.5	28.7	71	PCT-US02-30312-2166	Sequence 2166, Ap
27	81.5	28.7	71	PCT-US02-30412-2166	Sequence 2166, Ap
28	81.5	28.7	71	US-09-962-756-2166	Sequence 2166, Ap
29	81.5	28.7	71	US-10-253-471-2166	Sequence 2166, Ap
30	81.5	28.7	71	US-10-253-493-2166	Sequence 2166, Ap
31	81.5	28.7	73	PCT-US02-30312-2143	Sequence 2143, Ap
32	81.5	28.7	73	PCT-US02-30412-2143	Sequence 2143, Ap
33	81.5	28.7	73	US-09-962-756-2143	Sequence 2143, Ap
34	81.5	28.7	73	US-10-253-471-2143	Sequence 2143, Ap
35	81.5	28.7	73	US-10-253-493-2143	Sequence 2143, Ap
36	80.5	28.3	57	PCT-US02-30312-2147	Sequence 2147, Ap
37	80.5	28.3	57	PCT-US02-30312-2179	Sequence 2179, Ap
38	80.5	28.3	57	PCT-US02-30412-2179	Sequence 2179, Ap
39	80.5	28.3	57	PCT-US02-30412-2179	Sequence 2179, Ap
40	80.5	28.3	57	US-09-962-756-2147	Sequence 2147, Ap
41	80.5	28.3	57	US-09-962-756-2179	Sequence 2179, Ap
42	80.5	28.3	57	US-10-253-471-2147	Sequence 2147, Ap
43	80.5	28.3	57	US-10-253-471-2179	Sequence 2179, Ap
44	80.5	28.3	57	US-10-253-493-2179	Sequence 2179, Ap
45	80.5	28.3	57	US-10-253-493-2179	Sequence 2179, Ap

## ALIGNMENTS

related applications

RESULT 1  
PCT-US02-22821-108  
Sequence 108, Application PC/TUS0222821  
GENERAL INFORMATION:  
APPLICANT: HK Pharmaceuticals, Inc.  
APPLICANT: Koister, Hubert  
APPLICANT: Siddiqui, Sunaib  
TITLE OF INVENTION: Little, Daniel  
TITLE OF INVENTION: Capture Compounds, Collections Thereof  
TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex  
TITLE OF INVENTION: Compositions  
FILE REFERENCE: 24743-2305  
CURRENT APPLICATION NUMBER: PCT/US02/22821  
CURRENT FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: 60/306,019  
PRIOR FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/311,123  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: 60/363,433  
PRIOR FILING DATE: 2002-03-11  
NUMBER OF SEQ ID NOS: 149  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 108  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo Sapien  
PCT-US02-22821-108  
Query Match 100.0%; Score 284; DB 1; Length 49;



LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-717-811A-5

Query Match  
Best Local Similarity 100.0%; Score 284; DB 3; Length 98;  
Pred. No. 7.2e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLYQWLGAPVPPDPLEPRREVCELNPCDELADHIGFOEAYRRFYGPV 49  
Db 50 YLYQWLGAPVPPDPLEPRREVCELNPCDELADHIGFOEAYRRFYGPV 98

RESULT 5  
US-08-246-626-5  
Sequence 5, Application US/08246626  
GENERAL INFORMATION:  
APPLICANT: Hiroshi EGUCHI et al.  
TITLE OF INVENTION: Recombinant Human Osteocalcin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Displaywrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,626  
FILING DATE: 20-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/717,811  
FILING DATE: 19-Jun-1991  
APPLICATION NUMBER: US 08/131,932  
FILING DATE: 05-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 98 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:

ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-246-626-5

Query Match  
Best Local Similarity 100.0%; Score 284; DB 6; Length 98;  
Pred. No. 7.2e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLYQWLGAPVPPDPLEPRREVCELNPCDELADHIGFOEAYRRFYGPV 49  
Db 50 YLYQWLGAPVPPDPLEPRREVCELNPCDELADHIGFOEAYRRFYGPV 98

RESULT 6  
PCT-US01-08655-186  
Sequence 186, Application PC/TUS0108655  
GENERAL INFORMATION:  
APPLICANT: Hysq, Inc  
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 21272-065  
CURRENT APPLICATION NUMBER: PCT/US01/08655  
CURRENT FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: 09/522,929  
PRIOR FILING DATE: 2000-04-18  
PRIOR APPLICATION NUMBER: 09/770,160  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 09/668,317  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 09/695,783  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 09/728,628  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: 09/783,066  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 09/816,828  
PRIOR FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 584  
SOFTWARE: Custom  
SEQ ID NO 186  
LENGTH: 100  
TYPE: PRT  
ORGANISM: Homo sapiens

PCT-US01-08655-186

Query Match 100.0%; Score 284; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 7.4e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAIPVYPDPLERRREVCELNPCCDELADHIGFQEAIRRYGCV 49  
|||||  
DB 52 YLYQWLGAIPVYPDPLERRREVCELNPCCDELADHIGFQEAIRRYGCV 100

RESULT 7

PCT-US01-12010-3  
; Sequence 3, Application PC/UTUS0112010  
; GENERAL INFORMATION:  
; APPLICANT: Genesense Pharmaceuticals, Inc.  
; APPLICANT: Benlivena, Steven C.  
; APPLICANT: Chew, Anne  
; APPLICANT: Choi, Julie Y.  
; APPLICANT: Koshiy, Beena  
; APPLICANT: Rounds, Eileen  
; APPLICANT: Stephens, J. Claiborne  
; TITLE OF INVENTION: Haplotypes of the BGLAP Gene  
; FILE REFERENCE: MMH-0512PCT BGLAP  
; CURRENT APPLICATION NUMBER: PCT/US01/12010  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 60/195,840  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Homo sapien  
PCT-US01-12010-3

Query Match 100.0%; Score 284; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 7.4e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAIPVYPDPLERRREVCELNPCCDELADHIGFQEAIRRYGCV 49  
|||||  
DB 52 YLYQWLGAIPVYPDPLERRREVCELNPCCDELADHIGFQEAIRRYGCV 100

RESULT 8

US-10-143-899-23  
; Sequence 23, Application US/10143899  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PTZ47CIN  
; CURRENT APPLICATION NUMBER: US/10/143,899  
; CURRENT FILING DATE: 2002-05-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 127

; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (18)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (64)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-143-899-23

Query Match 100.0%; Score 284; DB 25; Length 127;  
Best Local Similarity 100.0%; Pred. No. 9.9e-28;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYQWLGAIPVYPDPLERRREVCELNPCCDELADHIGFQEAIRRYGCV 49  
|||||  
DB 79 YLYQWLGAIPVYPDPLERRREVCELNPCCDELADHIGFQEAIRRYGCV 127

RESULT 9

US-10-217-651-336  
; Sequence 336, Application US/10217651  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PZJ1CIN  
; CURRENT APPLICATION NUMBER: US/10/217,651  
; CURRENT FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: 09/760,491  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785

;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/244,617  
;; PRIOR FILING DATE: 2000-11-01  
;; PRIOR APPLICATION NUMBER: 60/225,268  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/226,368  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/251,856  
;; PRIOR FILING DATE: 2000-12-08  
;; PRIOR APPLICATION NUMBER: 60/251,868  
;; PRIOR FILING DATE: 2000-12-08  
;; PRIOR APPLICATION NUMBER: 60/229,344  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/234,997  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: 60/229,343  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,345  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,287  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,513  
;; PRIOR FILING DATE: 2000-09-05  
;; PRIOR APPLICATION NUMBER: 60/231,413  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/229,509  
;; PRIOR FILING DATE: 2000-09-05  
;; PRIOR APPLICATION NUMBER: 60/236,367  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/237,039  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,038  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/236,370  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/236,802  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,037  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,040  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/240,960  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/239,935  
;; PRIOR FILING DATE: 2000-10-13  
;; PRIOR APPLICATION NUMBER: 60/239,937  
;; PRIOR FILING DATE: 2000-10-13  
;; PRIOR APPLICATION NUMBER: 60/241,787  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/246,474  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/246,532  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/249,216  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,210  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/226,681  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/225,759  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/225,213  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/227,182  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/225,214  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/235,836  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: 60/230,438  
;; PRIOR FILING DATE: 2000-09-06  
;; PRIOR APPLICATION NUMBER: 60/215,135  
;; PRIOR FILING DATE: 2000-06-30

;; PRIOR APPLICATION NUMBER: 60/225,266  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/249,218  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,208  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,213  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,212  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,207  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,245  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,244  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,217  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,211  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,215  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,264  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,214  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,297  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/232,400  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/231,242  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/232,081  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/232,080  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/231,414  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/231,244  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/233,064  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/233,063  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,397  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,399  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,401  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/241,808  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,826  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,786  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,221  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/246,475  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/231,243  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/233,065

Query Match 100.0%; Score 284; DB 26; Length 127;  
Best Local Similarity 100.0%; Pred. No. 9,9e-28;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLYWMLGAPVPPDPLPRREVCELNDCDELADHIGFOEAYRRFYGPV 49  
DB 79 YLYWMLGAPVPPDPLPRREVCELNDCDELADHIGFOEAYRRFYGPV 127

```
RESULT 10
US-08-973-667-1
; Sequence 1, Application US/08973667
; GENERAL INFORMATION:
; APPLICANT: Sakakibara, Shunpei
; APPLICANT: Kimura, Terutoshi
; APPLICANT: Morimoto, Shigeto
; TITLE OF INVENTION: ANTI-GIU17-OSTEOCALCIN ANTIBODY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 North Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,667
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; APPLICATION NUMBER: PCT/JP97/01246
; FILING DATE: 10-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-88608
; FILING DATE: 10-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 9-43331
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 423-43
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 21
; OTHER INFORMATION: /product= "Gla"
US-08-973-667-1

Query Match          97.9%; Score 278; DB 13; Length 49;
Best Local Similarity 98.0%; Pred. No. 1.9e-27;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 North Glebe Rd. 8th floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,667
FILING DATE: 10-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP97/01246
FILING DATE: 10-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-88608
FILING DATE: 10-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9-43331
FILING DATE: 27-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 423-43
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Human
FEATURE:
NAME/KEY: Modified-site
LOCATION: 17
OTHER INFORMATION: /product= "Gla"
NAME/KEY: Modified-site
LOCATION: 21
OTHER INFORMATION: /product= "Gla"
US-08-973-667-2

Query Match          95.8%; Score 272; DB 13; Length 49;
Best Local Similarity 95.9%; Pred. No. 1.1e-26;
Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

: CITY: Arlington
: STATE: VA
: COUNTRY: USA
: ZIP: 22201-4741
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/973,667
: FILING DATE: 10-DEC-1997
: CLASSIFICATION: 435
:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP97/01246
: FILING DATE: 10-APR-1997
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-88608
: FILING DATE: 10-APR-1996
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.
: REGISTRATION NUMBER: 25,327
: TELEPHONE: 703-816-4000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-816-4000
: TELEFAX: 703-816-4100
:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Human
:
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 17
: OTHER INFORMATION: /product= "Gla"
:
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 21
: OTHER INFORMATION: /product= "Gla"
:
: US-08-973-667-3
:
: Query Match 95.8%; Score 272; DB 13; Length 49;
: Best Local Similarity 95.9%; Pred. No. 1.1e-26;
: Matches 47; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
:
: QY 1 YLYQWLGAPVYPDPLEPRREVCGLNDPCELDADHIGFOEAYRRFGPV 49
: 1 YLYQWLGAPVYPDPPLXPRRXVCXLNDPCELDADHIGFOEAYRRFGPV 49
:
: RESULT 13
: US-08-973-667-4
: Sequence 4, Application US/08973667
: GENERAL INFORMATION:
: APPLICANT: Sakakibara, Shunpei
: APPLICANT: Kimura, Terutoshi
: APPLICANT: Morimoto, Shigeto
: TITLE OF INVENTION: ANTI-GLI17-OSTEOCALCIN ANTIBODY
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHAYE P.C.
: STREET: 1100 North Glebe Rd. 8th floor
: CITY: Arlington
: STATE: VA
: COUNTRY: USA
: ZIP: 22201-4741

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/973,667
: FILING DATE: 10-DEC-1997
: CLASSIFICATION: 435
:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP97/01246
: FILING DATE: 10-APR-1997
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-88608
: FILING DATE: 10-APR-1996
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.
: REGISTRATION NUMBER: 25,327
: TELEPHONE: 703-816-4000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-816-4000
: TELEFAX: 703-816-4100
:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Human
:
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 17
: OTHER INFORMATION: /product= "Gla"
:
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 21
: OTHER INFORMATION: /product= "Gla"
:
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 24
: OTHER INFORMATION: /product= "Gla"
:
: US-08-973-667-4
:
: Query Match 93.7%; Score 266; DB 13; Length 49;
: Best Local Similarity 93.9%; Pred. No. 6.3e-26;
: Matches 46; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
:
: QY 1 YLYQWLGAPVYPDPLEPRREVCGLNDPCELDADHIGFOEAYRRFGPV 49
: 1 YLYQWLGAPVYPDPPLXPRRXVCXLNDPCELDADHIGFOEAYRRFGPV 49
:
: RESULT 14
: US-09-657-276-1093
: Sequence 1093, Application US/09657276
: GENERAL INFORMATION:
: APPLICANT: Conjuchem, Inc.
: APPLICANT: Bridon, Dominique
: APPLICANT: Ezrin, Alan
: APPLICANT: Milner, Peter
: APPLICANT: Holmes, Darren
: APPLICANT: Thibaudau, Karen
: TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
: TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
: TITLE OF INVENTION: COMPONENTS
: FILE REFERENCE: 2110
: CURRENT APPLICATION NUMBER: US/09/657,276
: CURRENT FILING DATE: 2000-09-07

```

PRIOR APPLICATION NUMBER: 60/134,406  
PRIOR FILING DATE: 1999-05-17  
PRIOR APPLICATION NUMBER: 60/153,406  
PRIOR FILING DATE: 1999-09-10  
PRIOR APPLICATION NUMBER: 60/159,783  
PRIOR FILING DATE: 1999-10-18  
NUMBER OF SEQ ID NOS: 1617  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1093  
LENGTH: 47  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-657-276-1093

Query Match: 89.1%; Score 253; DB 20; Length 47;  
Best Local Similarity 95.9%; Pred. No. 2,7e-24;  
Matches 47; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 1 YLYQWLGAAPVPPDLEPRREVCELNPDDELADHIGFQEAAYRRFGPV 49  
DB 1 YLYQWLGAAPVPPDLEPRR-VC-LNPDDELADHIGFQEAAYRRFGPV 47

## RESULT 15

US-09-036-085-5  
Sequence 5, Application US/09036085  
GENERAL INFORMATION:  
APPLICANT: Gary S. Stein et al.  
TITLE OF INVENTION: Gene Therapy Using Bone Marrow Transplants Transfected With T  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/036,085  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PROVISIONAL APPLICATION SERIAL NO. 60/039,839  
FILING DATE: March 6, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: UMW-007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-036-085-5

Query Match 84.5%; Score 240; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 APVPPDLEPRREVCELNPDDELADHIGFQEAAYRRFGPV 49  
|||||

DB 1 APVPPDLEPRREVCELNPDDELADHIGFQEAAYRRFGPV 42  
Search completed: December 4, 2002, 15:37:23  
Job time: 143 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 15:34:14 ; Search time 11 seconds

(without alignments)  
72.352 Million cell updates/sec

Title: US-09-462-931-2

Perfect score: 284  
Sequence: 1 YLYQWLGAVPYDPDLEPRR.....DELADHIGFQAVRYRYPV 49

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	26.8	13	10	US-09-822-485-29
2	76	26.8	13	10	US-09-801-968-21
3	76	26.8	136	10	US-09-858-349-2
4	60.5	21.3	677	10	US-09-815-242-10210
5	60	21.1	468	10	US-09-814-777A-2
6	60	21.1	468	10	US-09-814-777A-4
7	59	20.8	340	10	US-09-814-777A-15
8	59	20.8	384	10	US-09-814-777A-18
9	59	20.8	384	10	US-09-814-777A-20
10	59	20.8	470	10	US-09-814-777A-100
11	57	20.1	622	9	US-09-981-353-50
12	56	19.7	225	10	US-09-905-810-1
13	55.5	19.5	704	10	US-09-815-242-11921
14	53.5	18.7	677	10	US-09-815-242-11921
15	53	18.7	191	9	US-09-950-933A-66
16	53	18.7	216	10	US-09-924-358-47
17	53	18.7	458	10	US-09-815-242-11370
18	53	18.7	1550	10	US-09-995-542-8
19	53	18.7	2100	10	US-09-995-542-6

20	53	18.7	2144	10	US-09-858-194-2	Sequence 2, Appl
21	53	18.7	2146	10	US-09-995-542-5	Sequence 5, Appl
22	52.5	18.5	1005	10	US-09-802-127-5	Sequence 5, Appl
23	52	18.3	422	10	US-09-765-068-2	Sequence 2, Appl
24	52	18.3	459	10	US-09-797-039-5	Sequence 5, Appl
25	52	18.3	565	9	US-09-999-248-4	Sequence 4, Appl
26	51	18.0	103	10	US-09-841-132-501	Sequence 501, App
27	51	18.0	111	12	US-10-028-051-5	Sequence 5, Appl
28	51	18.0	409	10	US-09-815-242-11699	Sequence 11699, A
29	51	18.0	499	10	US-09-972-714-10	Sequence 6, Appl
30	50.5	17.8	1799	10	US-09-845-583-6	Sequence 6, Appl
31	50	17.6	1604	9	US-09-746-491-8	Sequence 8, Appl
32	49.5	17.4	1336	9	US-09-945-901-58	Sequence 58, Appl
33	49.5	17.4	1336	9	US-10-007-747-58	Sequence 58, Appl
34	49	17.3	43	9	US-10-041-406-6	Sequence 6, Appl
35	49	17.3	161	10	US-09-886-404-2	Sequence 2, Appl
36	49	17.3	162	10	US-09-925-299-1151	Sequence 1151, Ap
37	49	17.3	177	9	US-09-874-503-6	Sequence 6, Appl
38	49	17.3	177	12	US-10-006-867-156	Sequence 156, App
39	49	17.3	271	10	US-09-846-808-20	Sequence 20, Appl
40	49	17.3	371	10	US-09-828-302-15	Sequence 15, Appl
41	49	17.3	856	9	US-09-852-797-77	Sequence 77, Appl
42	49	17.3	856	10	US-09-853-161-77	Sequence 77, Appl
43	49	17.3	856	10	US-09-852-659A-77	Sequence 77, Appl
44	49	17.3	2257	10	US-09-767-479-10	Sequence 10, Appl
45	48.5	17.1	75	10	US-09-764-847-948	Sequence 948, App

#### ALIGNMENTS

RESULT 1  
US-09-822-485-29  
; Sequence 29, Application US/09822485  
; Patent No. US20020001825A1  
; GENERAL INFORMATION:  
; APPLICANT: Itch, No. US0020001825A1uyuki  
; TITLE OF INVENTION: No. US20020001825A1el Fibroblast Growth Factor-Like Polypeptid  
; FILE REFERENCE: 08035.0001-01000  
; CURRENT APPLICATION NUMBER: US/09/822,485  
; CURRENT FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:protein tag  
US-09-822-485-29

Query Match 26.8%; Score 76; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAVPYDPDLEPR 19  
|||||  
Db 1 GAVPYDPDLEPR 13

RESULT 2  
US-09-801-968-21  
; Sequence 21, Application US/09801968  
; Patent No. US20020082205A1  
; GENERAL INFORMATION:  
; APPLICANT: Itch, No. US20020082205A1uyuki  
; APPLICANT: Kavanaugh, W. Michael  
; TITLE OF INVENTION: HUMAN FG-23 GENE AND GENE EXPRESSION  
; FILE REFERENCE: PP-17150.001/201130.40901  
; CURRENT APPLICATION NUMBER: US/09/801,968  
; CURRENT FILING DATE: 2001-03-07  
; NUMBER OF SEQ ID NOS: 46

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: E tag
US-09-801-968-21

Query Match          26.8%; Score 76; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 GAPVPPDPLEPR 19
        |||
        1 GAPVPPDPLEPR 13

RESULT 3
US-09-858-349-2
; Sequence 2, Application US/09858349
; Patent No. US20020012909A1
; GENERAL INFORMATION:
; APPLICANT: PLASIN, Daniel
; TITLE OF INVENTION: SMALL FUNCTIONAL UNITS OF ANTIBODY HEAVY CHAIN VARIABLE REGIONS
; FILE REFERENCE: 87534-2800
; CURRENT APPLICATION NUMBER: US/09/858,349
; CURRENT FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 136
; TYPE: PRT
; ORGANISM: mouse hybridoma specific for H-2D + RGGRAFTV1 peptide
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (9)..(107)
; OTHER INFORMATION: variable
US-09-858-349-2

Query Match          26.8%; Score 76; DB 10; Length 136;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 GAPVPPDPLEPR 19
        |||
        122 GAPVPPDPLEPR 134

RESULT 4
US-09-815-242-10210
; Sequence 10210, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Truick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10210
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10210

Query Match          21.3%; Score 60.5; DB 10; Length 677;
Best Local Similarity 27.3%; Pred. No. 4.8;
Matches 12; Conservative 7; Mismatches 16; Indels 9; Gaps 1;

QY      1 YLYOMGAPVPPDPLEPRREVCELNPCODELADHIGFOEAYR 44
        |||
        250 YFYWLDAPIGYMGSEKML-----CDKRGDSVSFDEYWK 284

RESULT 5
US-09-814-777A-2
; Sequence 2, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING TH
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU PQ6457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2128)
; NAME/KEY: misc_feature
; LOCATION: (2315)..(3472)
US-09-814-777A-2

Query Match          21.1%; Score 60; DB 10; Length 468;
Best Local Similarity 39.5%; Pred. No. 3.6;
Matches 15; Conservative 3; Mismatches 14; Indels 6; Gaps 1;

QY      3 YOMLGAPVPPDPLEPRRE-----VCELNPDCELEAD 34
        |||
        368 YCTIGTGPGRFMRSLPPSPSLREGTEGLEPTADLMAD 405

RESULT 6
US-09-814-777A-4
; Sequence 4, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING TH
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU PQ6457
; PRIOR FILING DATE: 2000-03-24
```

```
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 468
; TYPE: PRN
; ORGANISM: Mouse
US-09-814-777A-4
```

```
Query Match
Best Local Similarity 21.1%; Score 60; DB 10; Length 468;
Matches 15; Conservative 3; Mismatches 14; Indels 6; Gaps 1;
```

```
Qy 3 YOWIGAPVPPDPLEPRRC-----VCLNPDDELAD 34
Db 368 YGTLGTGPPPGPLSPPEAPPLESAPPLGPAADLWAD 405
```

## RESULT 7

```
US-09-814-777A-15
; Sequence 15, Application US/09814777A
; Patent No. US20020142415A1
```

```
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU P06457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 340
; TYPE: PRN
; ORGANISM: Human
US-09-814-777A-15
```

```
Query Match
Best Local Similarity 20.8%; Score 59; DB 10; Length 340;
Matches 16; Conservative 0; Mismatches 16; Indels 6; Gaps 1;
```

```
Qy 3 YOWIGAPVPPDPLEPRRC-----ELNPDDELAD 34
Db 240 YGTLGTGPPPGPLSPPEAPPLESAPPLGPAADLWAD 277
```

## RESULT 8

```
US-09-814-777A-18
; Sequence 18, Application US/09814777A
; Patent No. US20020142415A1
```

```
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU P06457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 384
; TYPE: PRN
; ORGANISM: Human
```

```
; NAME/KEY: misc_feature
; LOCATION: (1)..(482)
; NAME/KEY: misc_feature
; LOCATION: (679)..(1919)
; OTHER INFORMATION: Exon 2
US-09-814-777A-18
```

```
Query Match
Best Local Similarity 20.8%; Score 59; DB 10; Length 384;
Matches 16; Conservative 0; Mismatches 16; Indels 6; Gaps 1;
```

```
Qy 3 YOWIGAPVPPDPLEPRRC-----ELNPDDELAD 34
Db 284 YGTLGTGPPPGPLSPPEAPPLESAPPLGPAADLWAD 321
```

## RESULT 9

```
US-09-814-777A-20
; Sequence 20, Application US/09814777A
; Patent No. US20020142415A1
```

```
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU P06457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 384
; TYPE: PRN
; ORGANISM: Human
US-09-814-777A-20
```

```
Query Match
Best Local Similarity 20.8%; Score 59; DB 10; Length 384;
Matches 16; Conservative 0; Mismatches 16; Indels 6; Gaps 1;
```

```
Qy 3 YOWIGAPVPPDPLEPRRC-----ELNPDDELAD 34
Db 284 YGTLGTGPPPGPLSPPEAPPLESAPPLGPAADLWAD 321
```

## RESULT 10

```
US-09-814-777A-100
; Sequence 100, Application US/09814777A
; Patent No. US20020142415A1
```

```
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU P06457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 100
; LENGTH: 470
; TYPE: PRN
; ORGANISM: Human
US-09-814-777A-100
```

```
Query Match
Best Local Similarity 20.8%; Score 59; DB 10; Length 470;
Matches 16; Conservative 0; Mismatches 16; Indels 6; Gaps 1;
```

```
Qy 3 YOWIGAPVPPDPLEPRRC-----ELNPDDELAD 34
Db 240 YGTLGTGPPPGPLSPPEAPPLESAPPLGPAADLWAD 277
```

```
RESULT 11
US-09-981-353-50
; Sequence 50, Application US/09981353
```

```
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 50
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 2264641CD1
; US-09-981-353-50
```

```
Query Match 20.1%; Score 57; DB 9; Length 632;
Best Local Similarity 36.8%; Pred. No. 12;
Matches 14; Conservative 5; Mismatches 15; Indels 4; Gaps 1;
```

```
QY 8 AYPVPDPLPRREVCELPDDELADHIGFOEAYRRF 45
DB 12 AYPVPVPVPMWRQFCFELHAQ---AAAVDFAHKFCRF 45
```

```
RESULT 12
US-09-905-810-1
; Sequence 1, Application US/09905810
; Patent No. US20020015703A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: RENNERT, Paul
; TITLE OF INVENTION: Antagonists of Tweak and of Tweak
; TITLE OF INVENTION: Receptor and their use to Treat Immunological Disorders
; FILE REFERENCE: A068 US
; CURRENT APPLICATION NUMBER: US/09/905,810
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: PCT/US00/01044
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/116,168
; PRIOR FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Murine
; US-09-905-810-1
```

```
Query Match 19.7%; Score 56; DB 10; Length 225;
Best Local Similarity 41.9%; Pred. No. 5.1;
Matches 18; Conservative 3; Mismatches 18; Indels 4; Gaps 2;
```

```
QY 2 LYQW-LGAPVPYDPL--EPREVCELPDDELADHIGFOE 40
DB 20 LGSWATLSAQEPSSQELTADRRREPPLNPTBESQDVYFFLE 62
```

```
RESULT 13
US-09-815-242-13925
; Sequence 13925, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
```

```
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13925
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Salmonella typhi
; US-09-815-242-13925
```

```
Query Match 19.5%; Score 55.5; DB 10; Length 704;
Best Local Similarity 27.3%; Pred. No. 21;
Matches 12; Conservative 6; Mismatches 17; Indels 9; Gaps 1;
```

```
QY 1 YLYQW-LGAPVPYDPLPRREVCELPDDELADHIGFOEAYRR 44
DB 277 YYYWLDAPIDAGTMSFKNL-----CDKRGDTTSPFETVKKK 311
```

```
RESULT 14
US-09-815-242-11921
; Sequence 11921, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11921
; LENGTH: 677
```

Wed Dec 4 16:22:10 2002

TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-815-242-11921

Query Match 18.8%; Score 53.5; DB 10; Length 677;  
Best Local Similarity 33.3%; Pred. No. 37;  
Matches 10; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 1 YLYQWLCAVPYPPDLEPRREVCLELNPDCD 30  
DB 249 YFYVWLDAPIGY--MASFKMLCARPELD 275

RESULT 15  
US-09-950-933A-66  
Sequence 66, Application US/09950933A  
Patent No. US20020166141A1  
GENERAL INFORMATION:  
APPLICANT: Simmons, Carl R.  
TITLE OF INVENTION: Antimicrobial Peptides and Methods of  
FILE OF INVENTION: Use  
FILE REFERENCE: 35718/238472  
CURRENT APPLICATION NUMBER: US/09/950,933A  
CURRENT FILING DATE: 2001-09-11  
PRIOR APPLICATION NUMBER: 60/232,569  
PRIOR FILING DATE: 2000-09-13  
NUMBER OF SEQ ID NOS: 99  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 66  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(48)  
OTHER INFORMATION: Xaa - Any Amino Acid  
US-09-950-933A-66

Query Match 18.7%; Score 53; DB 9; Length 191;  
Best Local Similarity 40.0%; Pred. No. 10;  
Matches 12; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

QY 8 APVYPPDPL--EPRREVCLELNPDCDELADH 35  
DB 109 SPVYPPPPVAPSPAPVYVKSNDICPLCDY 138

Search completed: December 4, 2002, 15:37:59  
Job time: 12 secs

